



Embeddable AntiVirus engine with high granularity

our understanding and dream

seak

Seak@antiy.net

- Challenges to AV Principles
- High granularity processing
- Embeddable AV Engine



Challenges to AV Principles

- AV is not simply a technological battlefield. The overall AV system takes in many logical and legal factors. There are also project planning factors which have some basic principles in common.
- These common principles can be objectively summarized from the AV practice, and then applied to guide the design of an AV engine and tools.
- In 1995 we summarized the basic common principles in 44 items, informally named AV dialectics.

- A computer virus is a kind of program in the final analysis
- The features of a computer virus are the only identifiers to classify the virus
- The crucial criterion in judging a program to be a virus should be its features or some characteristics of the content
- The only reason that feature code should be purged is if it is objectively or subjectively harmful
- Whether a certain program should be classified as a virus or not should be based on clear criteria
- **The clean up of a virus is the reverse of its infection**
- **User's rights to the AV software:**
 - Right to decide: Users can customize the functionality of the AV software instead of using the default configuration**
 - Right to know: Users should know what the AV software has done in the system**
 - Right to backup: Users should be provided with means to backup infected files**
- **Software should detect viruses inside packages and clean viruses without deleting the package if authorized**
- **Precaution principle: Virus monitoring should prevent the infected files from running and taking control of the system**

- With the development of both the application environment and virus techniques, many of our above stated points began to contradict each other
- The fundamental reason for these contradictions is the complication of information systems

- Item: The crucial criterion of a computer virus should be the feature code or some characteristics of the content
- Exception: CMD backdoor left by Code Red
- Question: Traditional AV technologies deal with “Yes or No” problems, where the only criterion is the content of the program. But under some circumstances, the boundary between harmful and harmless becomes vague.

- Item: Whether a certain program should be detected or not should be based on clear criteria
- Exception: psexec tool used in Worm.Dvldr .
- Question: The emergence of unwanted files is another puzzle in detection criterion. How far should AV software reach? What is the criterion? So far, many AV products include adware detection, is this reasonable or legal?

- Item: Detect viruses inside packages and clean viruses without deleting the package if authorized
- Exception: DIY worms (such as password worms), and worms using or saving in zip formats (such as some variation of netsky)
- Question: The basic assumption of traditional AV software is that a package file is normal file that may contain a virus. DIY worms are self-extracting packages. Some worms make many zipped backup copies on the disk which cannot be removed by AV software.

- Item: The only reason that feature code should be purged is if it is objectively or subjectively harmful
- Exception: Crisis caused by unofficial evaluation
- Problem: If one company detects some trivial files, other companies will follow suit in order to win higher marks in competitive evaluation. Is this worthwhile behavior? How can it be balanced with efficient and high-throughput virus detection?

- Item: The clean up of a virus is the reverse of its infection
- Case: Leftover backdoors leading to a worm returning
- Question: Is AV software responsible for recovering all the system modifications made by the virus? And how to deal with leaks? Is this work endless?

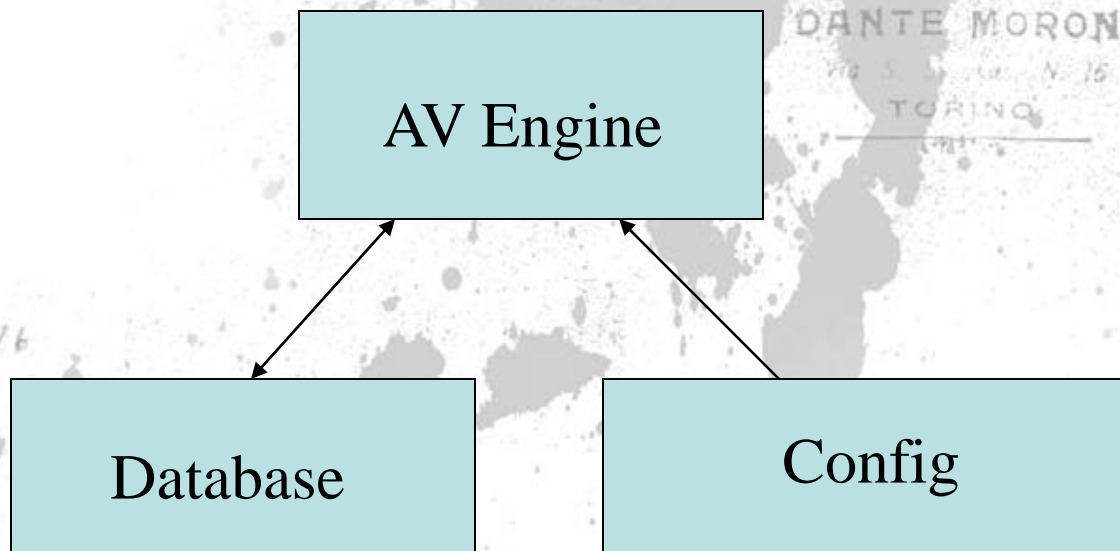
- Item: Virus monitoring should prevent infected files from running and taking control of the system
- Case: Arguments on file evaluation
- Question: Since it is difficult to detect unknown PE viruses, Trojans or backdoors, should the heuristic report based on behavior be acted on immediately?

- Item: User's rights to the AV software
- Case: scanning worms changed the image of victims
- Question: At first, viruses aimed simply to infect users' systems. Now, more often than not, they try to make infected systems further infect other systems. In such a case, can a virus be removed without the user's permission? What means are acceptable? Is this a technological question or legal question?

- None of these problems are too difficult to be solved technologically
- Some of them concern style and morals, however “Puzzling Criterion”, “Package Enigma”, and the “Responsibility Problem” are reactions to the traditional system and framework of the AV engine.
- We need more adaptive and reasonable engine frameworks instead of expediency in programming



High Granularity Processing



The three elements of an AV engine are the engine, database, and configuration. The engine relies on the database to detect, and the definitions in the configuration to work.

Before, we put much emphasis on the engine. Now, we need to pay more attention to the configuration to see what gains it has to offer us.

We also need to reevaluate the database – the maintenance of which is traditionally mechanical – to see whether the potential for creativity still exists.

	Type 1	Type 2	Type 3	Type 4
Number	✓	✓	✓	✓
Mod num	✓	✓	✓	✓
Virus name	✓	✓	✓	✓
First word of Feature code			✓	✓
Offset1+Sign 1			✓	✓
Offset2+Sign 2			✓	✓
File type flag				✓
Process arg	✓		✓	✓
Processing module name			✓	✓

- In working with a database, 95% of viruses are detected via records of type 3 and type 4 (featuring code detection). Detecting the remaining 5% of special viruses is done with records of type 1 and type 2 (independent module detection).
- Over 80% of viruses are processed via argument, and the remaining 20% via processing module.

- Object Control: what to detect
- Behavior Control: how to process
- Effectiveness Control: intensity of detection



- Flow control (Program)
- Debug Switch (Developer)
- INI control (User)

STUDIO FOTOGRAFICO
DANTE MORONI
Via S. ... N. 15
TORINO

17.3.1976

Art. 1111 Postal

FOTOGRAFIA
Umberto Bonivento
PIOLA
Via Promontore 4

Stornatissima
Contessa Pianta
voglio Austria

Atta unice amara
quadrone perche
si ricorda qualche
alta di me, che
nota off
lunato mit Pila



- Memory=Yes; check the memory
- Sectors=Yes; check the boot sector
- Files=Yes; check file system
- Packed=Yes; check packages
- Archives=Yes; check archives
- MailBases=Yes; check emails
- MailPlain=Yes; check encoded files
- FileMask=2; check the extended names
- UserMask= ?; user defined extension
- Exclude=No; Don't check customized extensions
- ExcludeMask= ; Don't check definition of extensions

- `InfectedAction=0`; remove viruses
- `InfectedCopy=No`; back up viruses
- `InfectedFolder=Infected`; back up folders
- `SuspiciousCopy=No`; back up suspicious files
- `SuspiciousFolder=Suspicious`; back up folders
- `Report=Yes`; generate logs
- `ReportFileName=Report.txt`; name of log file

- Warnings=Yes; Show warnings
- CodeAnalyzer=Yes; Open the code analyzer
- RedundantScan=Yes; Redundant scanning




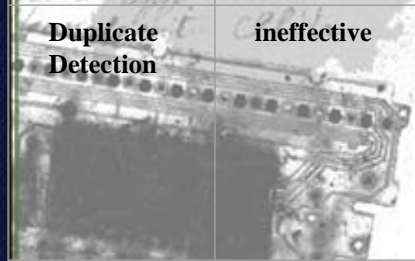
- In the traditional AV environment, this kind of granularity is enough for control, however problems occur when it comes to more complicated environments.



- Consider what different features the engine will have when working as AV software for a single computer VS working as one module in a mail server?
- I-Worm.Nimda.e is a infective worm. When processed locally, it should be regarded as a PE infected file, but for a mail server, it should simply be discarded.
- Win95.CIH is a infective virus. When detected, whether it is local or on mail server, it should be processed as an infected virus and the original file should be recovered.
- The essential difference is that Win95.CIH doesn't mail itself but rather is an executable program mailed by the user, while Nimda behaves contrarily.
- This situation requires different processes for different kinds of viruses in various environments, it is beyond the capacity of traditional engine control.

Application Case 2

- Network virus detection equipment contains several responding modules
- What policy should these responding modules work with?
- Some mail worms create addressees randomly, what will happen if sending creates a feedback loop?
- Some mail worms use bots to create addressees. What will happen if the worm starts sending duplicates?
- Email detection
- Duplicate email detection
- Feedback email detection
- Reset connection

	SMTP detection				POP3 detection			
	Faked Recipient	True Recipient	True Sender	Faked Sender	Faked Recipient	True Recipient	True Sender	False Sender
 <p>Feedback Detection</p>			effective	ineffective			effective	ineffective
 <p>Duplicate Detection</p>	ineffective	effective			effective	ineffective		

- Integration with networking equipment is an effective response.
- See: OPSEC, TOPSEC
- Different processing for scanning worms and mail worms.
- It is simple to scan worm infection IP nodes. But if we do the same to email worms, they may send the same email over and over again causing DoS
- We should check whether there is a proxy server on the network

- New demand goes beyond the capacity of the traditional engine
- How can we solve this problem?

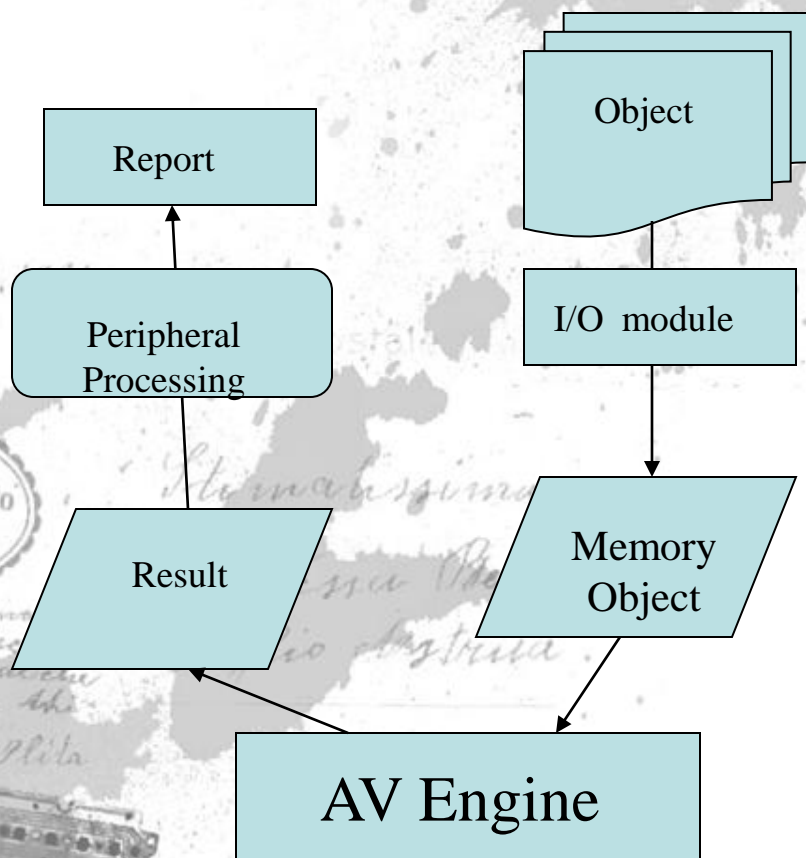




Embeddable AV engine with high granularity

- The trending movement of network security products implies that virus filtering mechanisms will extend to equipment at different levels
- The above discussion shows the need for the AV engine to adapt to more complicated environments
- Embedded equipment or AV engines in other environments are designed for high granularity

Application Form	Details
AV module in Firewall	<ul style="list-style-type: none">Construct linear speed virus filtering module for package filtering firewall with a network engine.Construct file stream virus filter for app proxies, transparent proxies or a stream filtering firewall with a file engine.
AV module in router	Add virus filtering ability to routing equipment with high speed package level scanning
AV module in switch	Add virus filtering ability to switching equipment with high speed packet scanning.
Virus detecting plug-ins in IDS	Extend the network engine to provide the IDS with network virus detection ability
AV module in GAP device	Extend GAP equipment with virus filtering ability
Virus protection in mail system	Embed virus detection ability into mail servers
Independent AV software	User need only to program against an API to develop their own AV software Antiy Labs www.antiy.net



```
/*scanning parameter structure */
typedef struct _AVLF_SDK_SCAN_PARA
{
char * pBuffer;                /* pointer to buffer */
unsigned long ulSize;          /*size of the buffer */
const char * pDescription;     /* description information */
int bUnpack;                   /* whether to unpack*/
int bKill;                      /* whether to kill the virus */
int bKilled;                    /* whether virus was killed successfully*/
} AVLF_SDK_SCAN_PARA,*PAVLF_SDK_SCAN_PARA;

/* set the receiver */
AVLEACHSDK_API int AVLF_SDK_SetReciver(IReportReciver *pReciver);

/*scanning: return 0 if no virus detected, return 1 if virus found, detailed information is received by
the receiver class*/
AVLEACHSDK_API int AVLF_SDK_Scan(PAVLF_SDK_SCAN_PARA pParamter);
```


- Modern AV engines have evolved from branched engines led by module-based format recognition to recursive engines
- In a recursive engine, scanned objects could have multiple flags, which can be detected by corresponding modules
- McAfee's bug in detecting SFX
 - `archbomb.zip`

ArcBomb.ZIP
ArcBomb
42 KB

lib	book	chapter	doc	page
lib 0.zip	book 0.zip	chapter 0.zip	doc 0.zip	page 0.zip
lib 1.zip	book 1.zip	chapter 1.zip	doc 1.zip	page 1.zip
lib 2.zip	book 2.zip	chapter 2.zip	doc 2.zip	page 2.zip
lib 3.zip	book 3.zip	chapter 3.zip	doc 3.zip	page 3.zip
lib 4.zip	book 4.zip	chapter 4.zip	doc 4.zip	page 4.zip
lib 5.zip	book 5.zip	chapter 5.zip	doc 5.zip	page 5.zip
lib 6.zip	book 6.zip	chapter 6.zip	doc 6.zip	page 6.zip
lib 7.zip	book 7.zip	chapter 7.zip	doc 7.zip	page 7.zip
lib 8.zip	book 8.zip	chapter 8.zip	doc 8.zip	page 8.zip
lib 9.zip	book 9.zip	chapter 9.zip	doc 9.zip	page 9.zip
lib a.zip	book a.zip	chapter a.zip	doc a.zip	page a.zip
lib b.zip	book b.zip	chapter b.zip	doc b.zip	page b.zip
lib c.zip	book c.zip	chapter c.zip	doc c.zip	page c.zip
lib d.zip	book d.zip	chapter d.zip	doc d.zip	page d.zip
lib e.zip	book e.zip	chapter e.zip	doc e.zip	page e.zip
lib f.zip	book f.zip	chapter f.zip	doc f.zip	page f.zip

35 KB 29 KB 32 KB 162 KB 4,071 KB

i0.dll 4,194,304 KB 2000-3-28 18:03

```
00000000h: AA AA AA AA AA AA AA AA AA AA AA AA AA AA AA ;
00000010h: AA AA AA AA AA AA AA AA AA AA AA AA AA AA AA ;
00000020h: AA AA AA AA AA AA AA AA AA AA AA AA AA AA AA ;
00000030h: AA AA AA AA AA AA AA AA AA AA AA AA AA AA AA ;
00000040h: AA AA AA AA AA AA AA AA AA AA AA AA AA AA AA ;
00000050h: AA AA AA AA AA AA AA AA AA AA AA AA AA AA AA ;
00000060h: AA AA AA AA AA AA AA AA AA AA AA AA AA AA AA ;
00000070h: AA AA AA AA AA AA AA AA AA AA AA AA AA AA AA ;
00000080h: AA AA AA AA AA AA AA AA AA AA AA AA AA AA AA ;
```

```
00000000 50 48 03 04 48 00 00 00 56 88 00 00 09 00 00 00 6C 69 .....V.....li
00000010 DC C8 F9 09 00 00 56 88 00 00 09 00 00 00 6C 69 .....V.....li
00000020 62 20 33 2E 7A 69 70 ED DD 57 50 53 DB 1A 07 F0 B 3.zip..WFS....
00000030 48 11 38 74 95 43 00 B1 51 22 48 95 1A C4 A0 DC H.St.C.G.H....
00000040 23 ED 09 48 95 62 28 5A 8A E8 0D 2D 14 2E 47 E9 #.H.b(Q...+G.
00000050 10 01 29 46 94 8E 28 51 44 7A B3 25 A0 20 08 A2 .J.F..(Qzr.%...
00000060 20 52 54 04 A4 B7 43 09 88 1C BD 8E F7 C6 87 33 RT...C.....3
00000070 37 33 79 FD F6 7E 58 7B CD DA FF F9 5E 7F B3 F7 73y...X{.....^..
00000080 2A 16 C6 AC 6C 5B 10 2C 08 4E 44 DB 9D B0 DD 54 *...[,]..ND...T
00000090 D7 F1 0D 92 9C 08 C4 46 3C 02 F1 1B 02 81 C0 FA .....F.....<
000000A0 F9 79 EF 50 51 08 39 71 6A B2 CF 5B D3 7A 75 1B .y.PQ.9qj..[.zu.
000000B0 C7 8C 14 25 6A 2B 97 90 44 AA B4 04 3A 39 D4 3D ...*}.D...:9..
000000C0 B7 E3 4F C4 40 A4 E0 AE 0A 11 93 29 A1 AC EA ED ..0.0.....
000000D0 96 81 E4 46 79 D3 08 49 C7 4D FF E6 EF 4E 2E 71 ...Fy..I.M...K.q
000000E0 A2 4A 25 79 3F AE F4 49 15 AC BA 34 4A 6A 25 9C .J%y?..I...4J%>
000000F0 6E 75 6E C3 67 E2 9C 6F D0 8A 17 B0 FB 1E 46 D6 nun.g.o.....F.
00000100 D7 C9 0E 73 A5 99 01 B4 B5 17 5C B5 D7 28 89 A6 ...s.....\{..(
00000110 4D 2E 89 DD 17 12 D7 90 4F 7A 3A F4 30 6C 86 5C M.....Dzr.01.\
00000120 87 04 76 B2 1D 24 19 9A 61 36 90 CF DC 17 95 7D .v..$.a6.....)
00000130 E4 7E 23 7D 83 79 5C EE 8E 5B B6 F8 FD 81 8A 26 ..#)y{.....6
00000140 F9 22 29 26 D9 BE C2 AD 96 52 61 05 B8 0B 75 1F *)%.Ra...u.
00000150 F7 39 4A 64 E7 16 1A 04 4B 78 68 1C 4E 4F 5E 7B .9d.....Kxh.M0[
00000160 30 74 CE 8B 1C 55 DB 94 77 5B 10 5D BF 66 A4 97 0c...U...w[.]h...
00000170 23 D3 68 67 5A 93 13 54 A0 45 CB 8E 5C 14 AA 68 #.hgZ..T.E...h
00000180 6A 9A AA C6 2C B6 86 4E D4 D7 5C EB D7 5E DE EF .....N...A.....
00000190 9B 42 9E 3F A0 7F FD EB 42 0D CF 99 ED 64 B3 EC ..B?...B...d...
000001A0 58 73 4C 7E 2E CD 2C 28 95 B5 83 D3 3B D9 63 5B Xsl...B...c[
000001B0 B8 AD CA 5D 9F 8E EE E9 0F 3A 4B 47 1C AD 58 91 .....].....K%>.X.
000001C0 FC 5A 38 72 C5 CA 26 F3 10 13 87 5B B1 CA 21 B6 .28r...&...[...!
000001D0 1E 0A 16 9E 1A 7B DE 18 6B 57 3D 35 B8 36 E9 AA .....{..k%5.6...
000001E0 F4 71 96 4F 8C F5 D3 B5 82 D4 C0 D4 1A 0E AF 89 .q.0.....
000001F0 2F 89 C2 D5 41 4D 9B FB 1F F9 0A 0F 74 14 B0 B4 /...AM.....E...
00000200 F1 BD 2C 6B 46 9A 97 E1 74 EC 72 37 52 85 0D 0B ...kF...t.z7R...
00000210 29 E1 7B F9 9A BD AE 6F 9E 53 B0 F5 E8 71 F4 1F ).(....o.S...q...
00000220 77 19 42 C5 7E DE B2 C2 96 4C B2 A9 29 B8 37 26 w.B...L...).7%
00000230 17 E8 85 CD C2 39 A0 8A 4B D1 AE 72 FB D1 CA C1 .....9...K...E...
00000240 B4 FB 2E 1F FB 89 E7 C3 83 F6 98 A9 28 78 1D CF .....&.....(x...
00000250 BF 3A 2C 9D A1 10 B9 B5 5C 6D E7 F5 E6 FA 30 6B .4.....v...m...Ok
00000260 A2 83 92 DA E2 57 E7 D1 55 77 63 AB 3F 57 A4 BB .....W...Dve...?U...
00000270 4B E4 6D 26 33 D8 59 4A 9C 1D 65 70 B9 36 #.me3.VJ...ep...6
00000280 BD B3 5F 9B 1D A7 4D 51 6E 64 4F F0 D5 6D B1 EA .....MQnd0...m...
00000290 71 D1 DF DD 90 76 A8 6A A6 4A D1 8F BA 78 37 6B q...v.j.J...x7k
000002A0 30 7E 35 1E 93 1B 7E 3A F6 CB 97 FC F6 2B 7D 7F 0-S...:.....+).
000002B0 A5 A1 D0 A3 D8 4D 7D 3C 2E C7 25 17 8C 3A 94 64 .....M)<...%.d
000002C0 76 6B CE 18 7B 4B 35 9C E8 2B 4F 67 75 8B 4D 09 vk...K5...+Ogu.M.
000002D0 89 6B A0 F9 BD 4F C1 8B A9 DF 69 E3 92 A3 56 2D .k.....0...i...V-
000002E0 9A D5 4E F4 1D 0D 93 1C B6 D5 D0 51 C5 38 D3 DC .K.....0.8...
000002F0 8E 0C 7E 50 AA 63 95 96 C7 47 DB ED FC AF 53 AD .x.P.c...G...S...
00000300 89 8F 15 5F 72 46 79 06 FT FA 5A 26 2D 34 5F 97 28 0F C9 D3 BA C5 AF 6F C8 83 9B 34 BE 8A 53
00000310 28 0F C9 D3 BA C5 AF 6F C8 83 9B 34 BE 8A 53 AC (...>...o...4...S...
00000320 CF 3C 51 BB 40 74 39 30 A0 8B AB 9E 48 2C 0A 5E <0.q.t90...H...^
00000330 41 90 71 73 BA 76 B2 7E D7 AA CD 5E 34 46 7F CE A.q.s.v...^4F...
00000340 A1 F9 A7 3E 78 C9 B3 62 FE F6 ED C5 59 94 EA 96 ...>...b...Y...
00000350 A9 C2 85 BB C5 48 3F AA C9 D5 97 B6 33 7C 6A E2 .....H?...3]...
00000360 A9 B8 D2 8E 5C 92 38 5B 71 7D 7E B2 64 89 2D 73 .....8[q]v.d...s
00000370 88 AD 3E AF 6C 2F 76 6B F3 DA BE CC A3 5D 9A E8 h...l/vk.....]...
00000380 B0 2D 92 41 E2 74 91 3E 17 27 6F 3E 1E 27 6F 3E 1E .A...t...>D...E...
00000390 D2 52 2E 50 4D C4 15 39 7E 21 D4 48 24 FE 84 0E .R.PB...9(1.H...E...
000003A0 77 85 98 3D 39 DB A4 ED 59 6F 68 95 05 B2 FC D3 w...=9...Yoh...
000003B0 E3 A6 24 8A 63 DD BE 1E F9 3E 99 6E EA 45 8E 14 .<.<.<...>n.E...
000003C0 A3 93 BE 15 78 72 3E 41 E4 B6 1A BA B7 BF A7 CB ...x%>A...
000003D0 6D 80 B6 63 2A 61 7F B8 1B 67 54 44 0C F6 D9 ...c'a...>.tD...
000003E0 97 96 2A 6E E2 7C 77 DD FB 37 3E 1E 42 25 DD 9D .n.lw...7..B4...
000003F0 63 F0 AA 1E 6A C9 CB FB F6 99 57 6C 91 44 D7 9E c...j...W...D...
00000400 7D AA 0E B7 57 D7 8F 7B 15 5F 36 F4 3C A5 AB DE ]...W...(.6.<...<
```



14 00 02 00 08 00 1B

sign1
Offset: 4h
Length: 7h



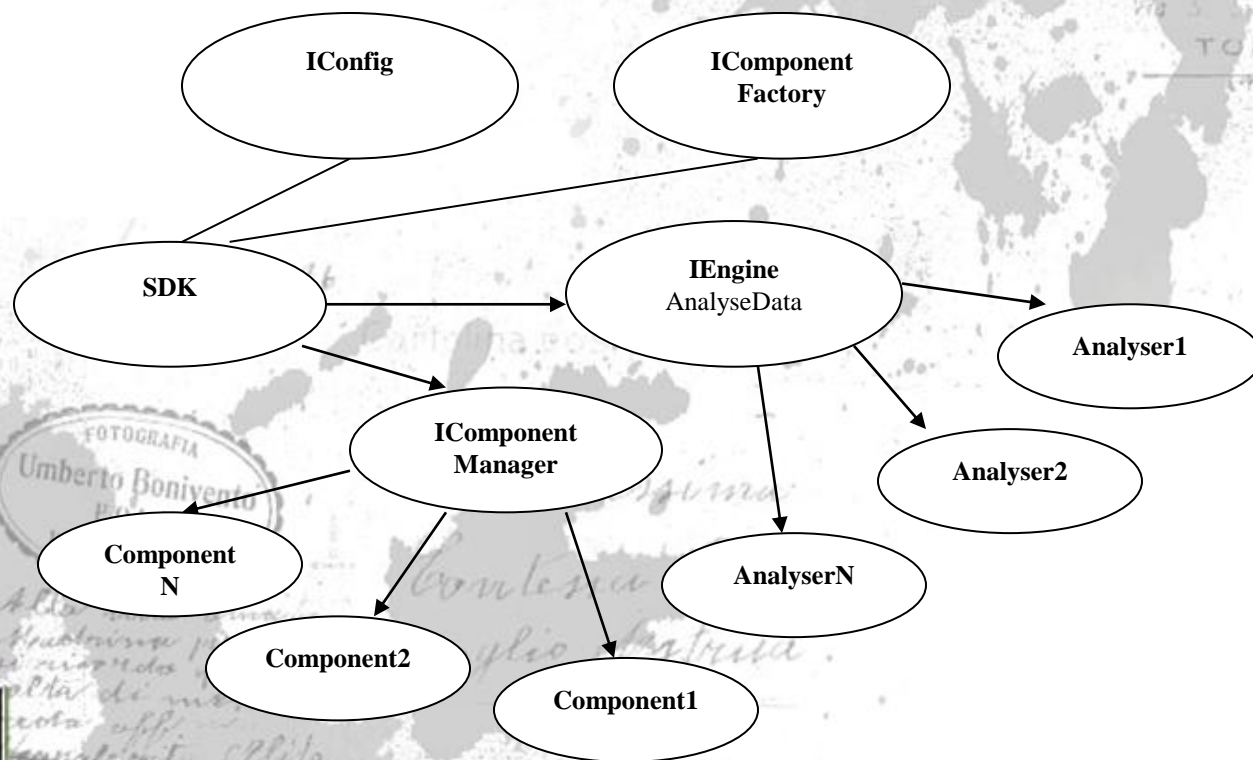
89 8F 15 5F 77 46 79 06 FT FA 5A 26 2D 34 5F 97 28 0F C9 D3 BA C5 AF 6F C8 83 9B 34 BE 8A 53
AC FC 31 5B 40 74 39 30 A0 8B AB 9E 48 2C 0A 5E 41 90 71 73 BA 76 B9 7E D7 AA CD 5E 34 46
7F CE A1 F9 A7 3E 78 C9 B3 62 FE F6 ED C5 59 94 EE 96 A9 C2 85 BB C5 48 3F AA C9 D5 97 B6 33
7C 6A E2 A9 B8 D2 8B 5C 99 38 5B 71 7D 76 E2 64 89 9D 73 68 AD 3E AF 6C 2F 76 6B F3 0A EB CC
A3 5D 8A F8 B0 20 92 41 EA 74 81 3E 17 27 6F 8C 85 72 03 09 D9 52 9F 50 40 C4 15 39 7B 21 D4
48 24 FE 84 08 77 85 98 3D 39 DB A4 ED 59 6F 68 95 05 B2 FC D3 E3 A5 24 8A 63 D0 BD 1E F9 3E
99 6E EA 45 8E 14 A3 93 BE 15 78 72 3E 41 E4 B6 1A BA B7 BF A7 CB 6D 80 B6 63 2A 61 7F FB 3B
1B 87 54 44 0C FE D9 97 96 2A 6E E2 7C 77 DD B5 37 BE 1A 42 25 DD 9D

Sign 2
Offset: 300h
Length: F0h



A Zip which is also a binary stream could be detected by binary engine instead of what would be done in the traditional branched engine – being passed to archive extracting module by the

Antiy L format recognition module.



1. Analyzers are parallel in structure, none are prerequisites.
2. Results from the analyzers can provide different priority ratings, with viruses listed as the highest and files needing further processing as the lowest
3. In principle, analyzers work serially, with higher priority results being forwarded.

- Working environment could be the x86 architecture, or other architectures like PPC
- Modules written in x86 assembly language are a barrier to porting to other architectures.



- What are the essential requirements for high granularity?
- Virus processing in different environments cannot only rely on detecting the infection feature but also the “specialty” of the virus.
- The granularity of control needs to reach the individual virus, the database needs to provide more information.
- Virus processing will be done with information from the database about the virus specialty.

- Flow control (Program)
 - Debug Switch (Developer)
 - INI control (User)
- Flow control (Program)
 - Virus attribute
 - Debug Switch (Developer)
 - Stencil (Condition)
 - INI control (User)

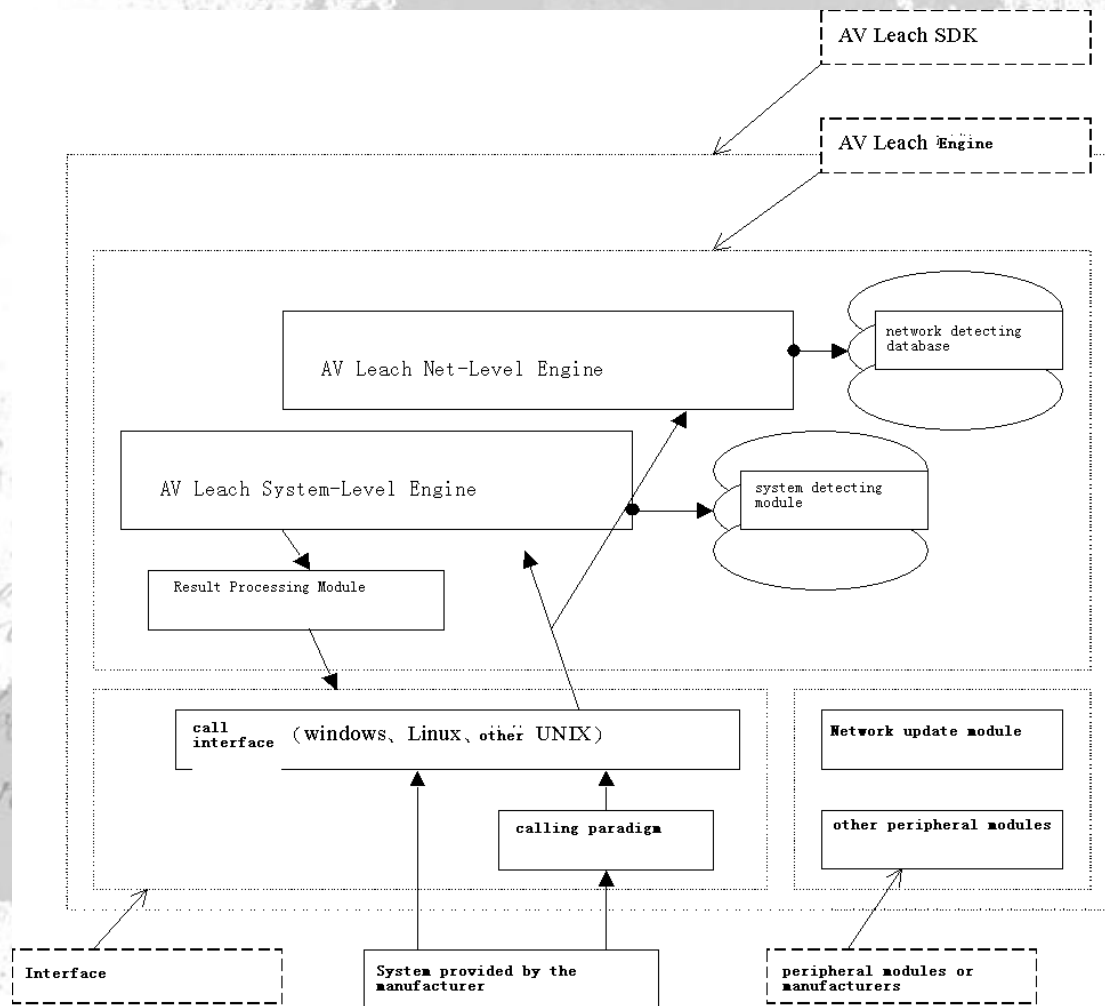
```
struct vxdb
{
char name[255];
char fword[4];
char offset1[4];
char crc1[8];
char offset2[4];
char crc2[8];
...
};
```

```
struct tgvxdb
{
char name[255];
char fword[4];
char offset1[4];
char crc1[8];
char offset2[4];
char crc2[8];
...
int vxattribute ;
};
```


- Perfect reverse engineering is the end goal.
- The High Granularity Engine ends the era in which the AV company does not need to analyze the virus.



- Clean com tail
 - Clean com head
 - Clean exe tail
 - Clean ne tail
 - Clean pe tail
 - Remove file
 - Copy data block
 - Move data block
 - Insert data block
 - Modify data block
 - Delete data block
 - Fill in data block
 - Truncate data tail
 - Truncate data head
- On the left is the cleaning parameter set which is widely accepted by many companies.
 - We need the same detailed processing script for non-infective viruses
 - Is this work endless?



- AV principles are not invariable. Instead, they are evolving dynamic principles. They require not only summarizing but also supplementing and replacing.
- We believe in our understanding and we persist in our dream.
- Thank you!