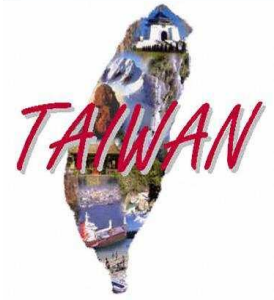
財團法人國家實驗研究院
國家高速網路與計算中心
National Center for High-Performance Computing

Improve Security-Events-Center to the Cloud Platform

Building ICAS with Hadoop and HBase



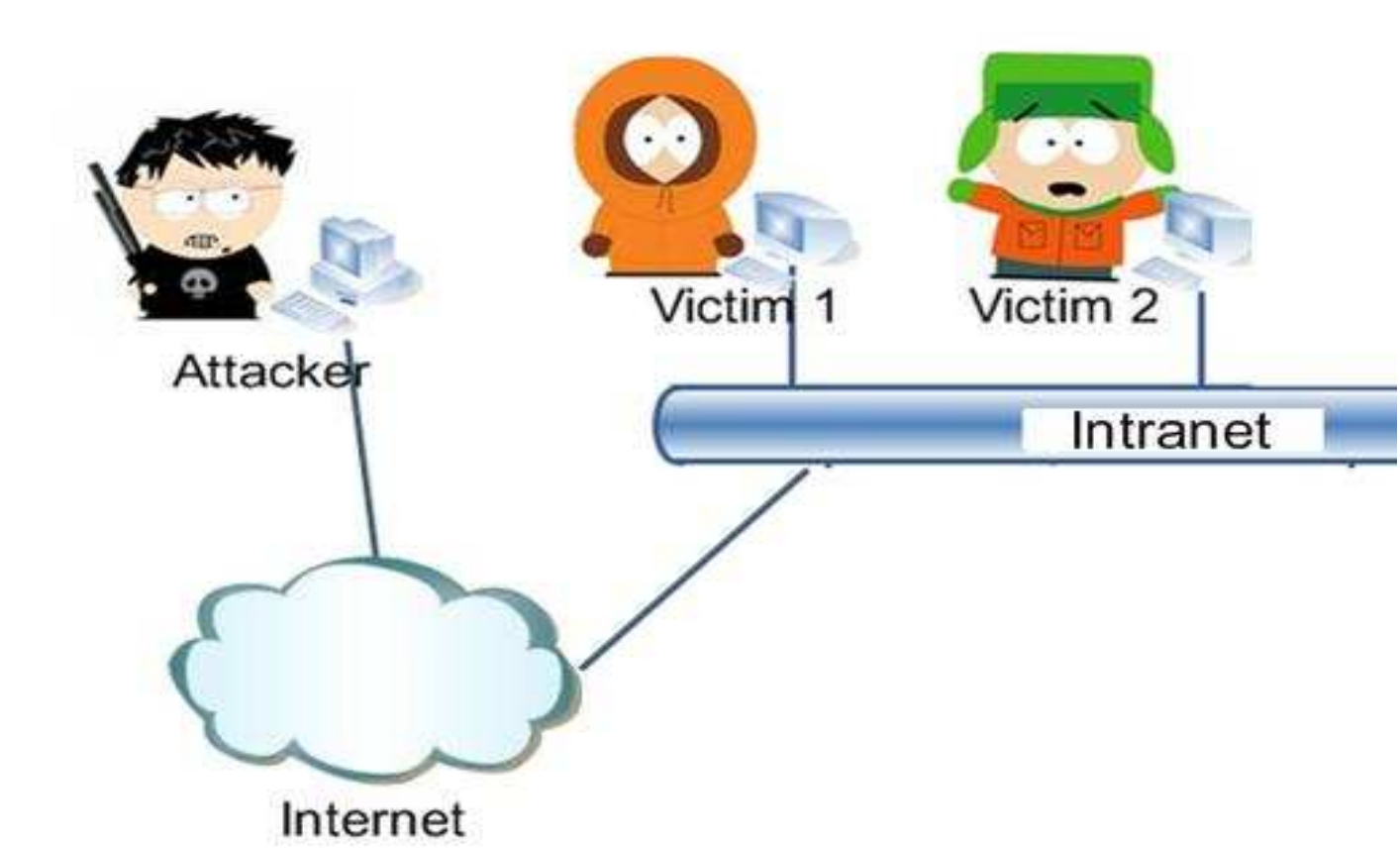
Wei-Yu Chen, Yao-Tsung Wang
National Center for High-Performance Computing, Taiwan
{waue,jazz}@nchc.org.tw

DATE: 05/08/2009

Outline

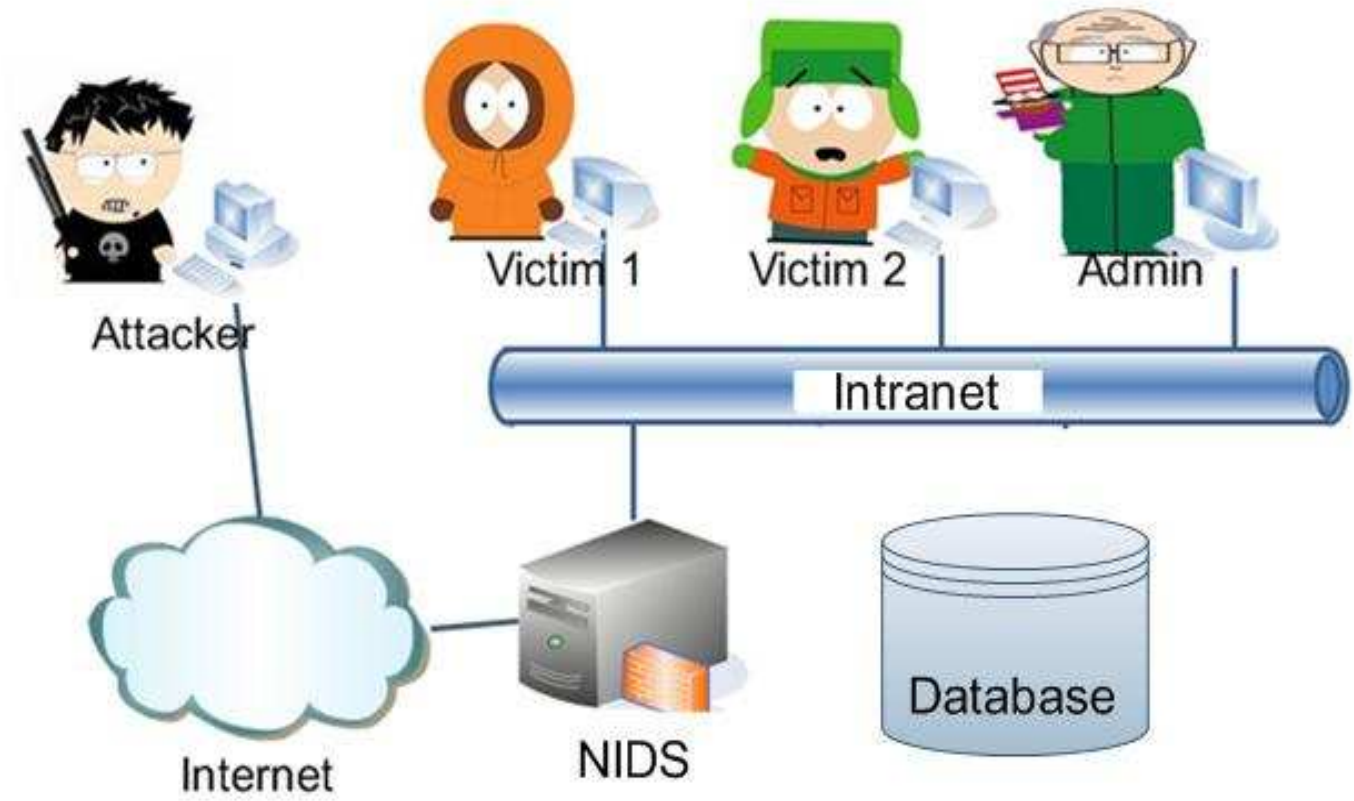
- The Background Story
- Our Idea and Methods
- Experiment Results
- Conclusions
- Future Works

Personal Security



Install Some Security Software

Internet Security



Network IDS Interface

Basic Analysis and Security Engine (BASE): Query Results - Mozilla

File Edit View Go Bookmarks Tools Window Help

Basic Analysis and Security Engine (BASE)

Home | Search | AG Maintenance

[Back]

Added 0 alert(s) to the Alert cache

Queried DB on : Thu October 14, 2004 22:04:44

Meta Criteria	*any*
IP Criteria	*any*
TCP Criteria	*any*
Payload Criteria	*any*

Summary Statistics

- **Sensors**
- **Unique Alerts** (**classifications**)
- Unique addresses: **source** | **destination**
- **Unique IP links**
- **Source** Port: **TCP** | **UDP**
- **Destination** Port: **TCP** | **UDP**
- **Time profile** of alerts

Displaying alerts 1-50 of 81 total

ID	< Signature >	< Timestamp >	< Source Address >	< Dest. Address >	< Layer 4 Proto >
#0-(1-84)	[snort] NETBIOS SMB IPC$ share unicode access	2004-10-08 11:25:41	192.168.1.100:1613	192.168.1.4:139	TCP
#1-(1-83)	[snort] NETBIOS SMB IPC$ share unicode access	2004-10-08 11:25:31	192.168.1.100:1608	192.168.1.4:139	TCP
#2-(1-82)	[snort] NETBIOS SMB IPC$ share unicode access	2004-10-08 11:25:05	192.168.1.100:1601	192.168.1.4:139	TCP
#3-(1-80)	[snort] (http_inspect) OVERSIZE CHUNK ENCODING	2004-10-04 22:25:41	192.168.1.4:42164	67.19.245.228:80	TCP
#4-(1-81)	[snort] (http_inspect) OVERSIZE CHUNK ENCODING	2004-10-04 22:25:41	192.168.1.4:42163	67.19.245.228:80	TCP

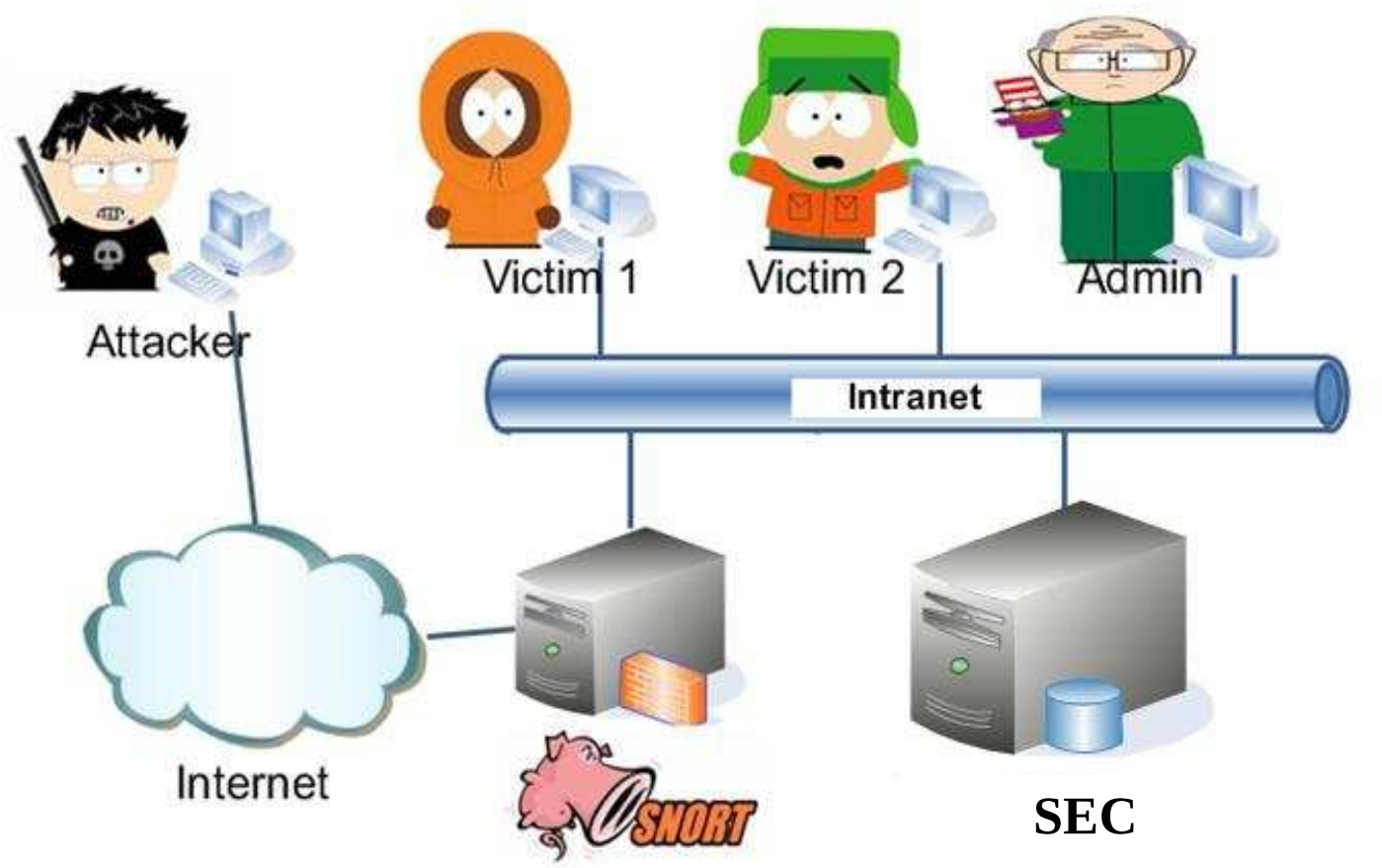
These Events are MIS's Nightmare !!!!

1. **Difficult to realize the overall accidents**
2. **Ignoring the crucial information easily !!!**

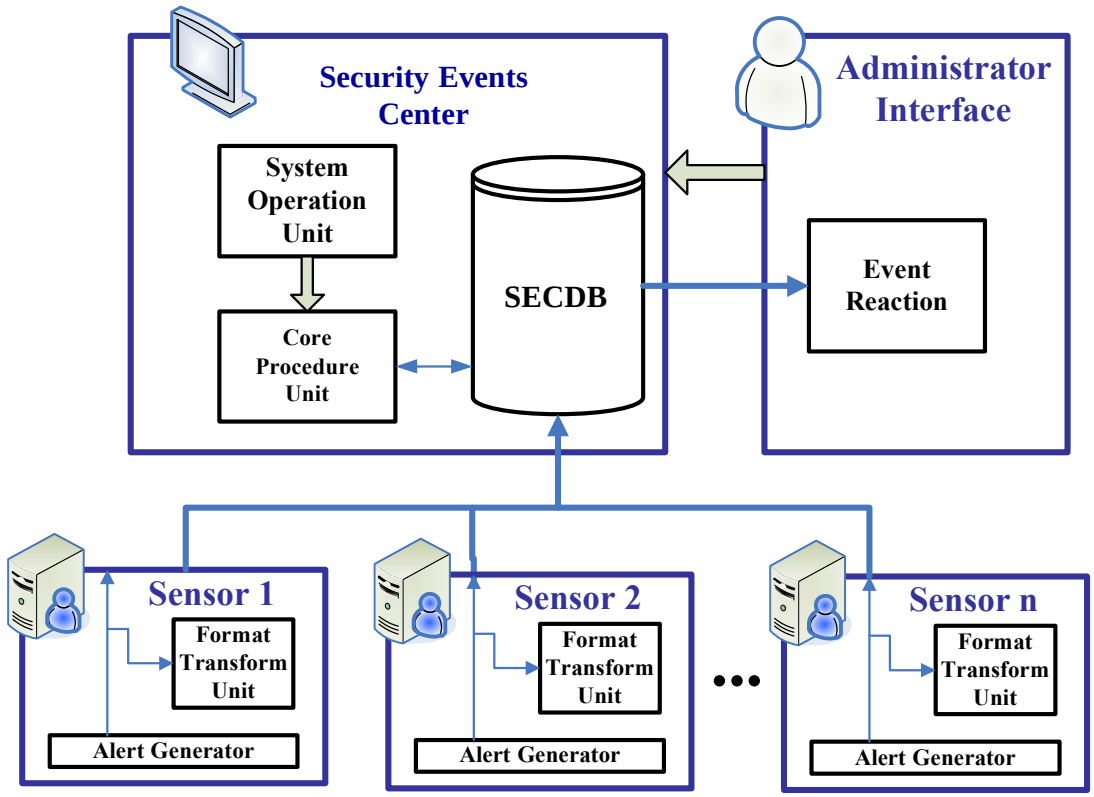
The Security Events Center

- A platform whose purpose is to provide detection and reaction services to security incidents.
- Main functions
  - Collects all information from both security and non-security products
  - Carries out the unified automatic event evaluation to tell if they are complying with the policy.

SEC Overview



The SEC Component

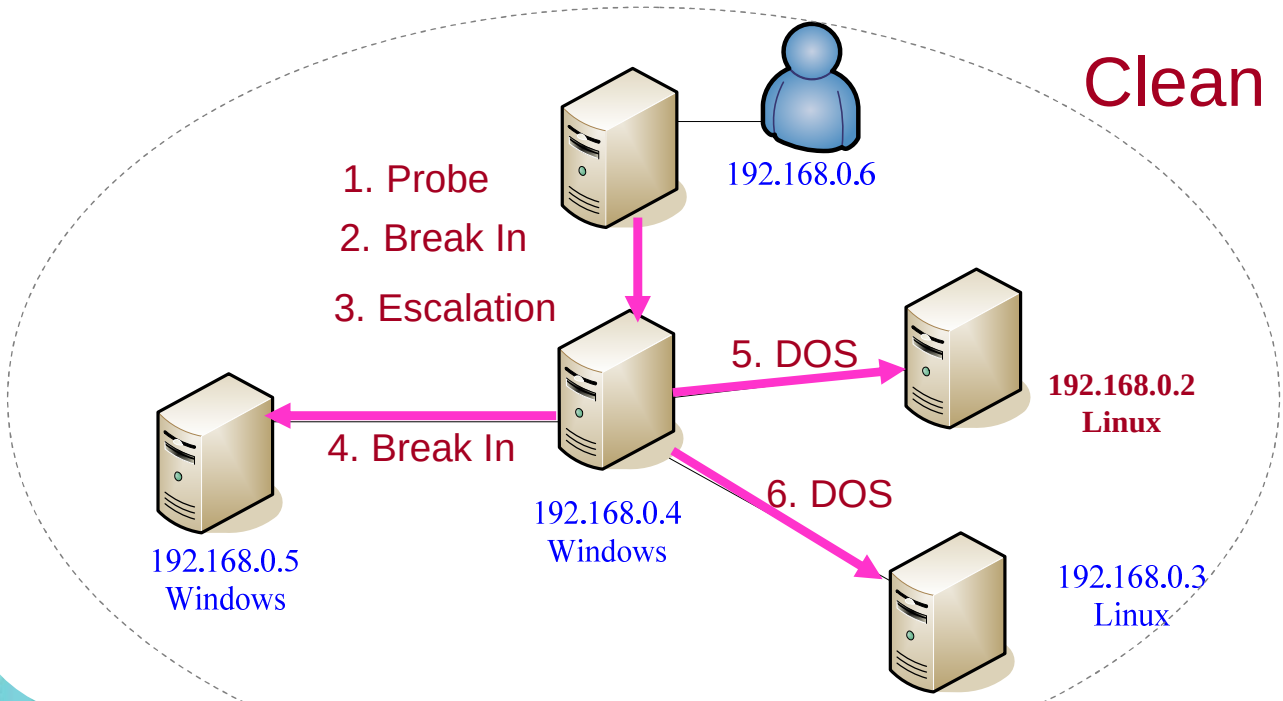


Alert Merge Example

Destination IP	Attack Signature	Source IP	Destination Port	Source Port	Packet Protocol	Timestamp
Host_1	Trojan	Sip1	80	4077	tcp	T1
Host_1	Trojan	Sip2	80	4077	tcp	T2
Host_1	Trojan	Sip1	443	5002	tcp	T3
Host_2	Trojan	Sip1	443	5002	tcp	T4
Host_3	D.D.O.S	Sip3	53	6007	udp	T5
Host_3	D.D.O.S	Sip4	53	6008	tcp	T5
Host_3	D.D.O.S	Sip5	53	6007	udp	T5
Host_3	D.D.O.S	Sip6	53	6008	tcp	T5

Key		Values				
Host_1	Trojan	Sip1,Sip2	80,443	4077,5002	tcp	T1,T2,T3
Host_2	Trojan	Sip1	443	5002	tcp	T4
Host_3	D.D.O.S.	Sip3,Sip4,Sip5 ,Sip6	53	6007,6008	tcp, udp	T5

Experiment : Scenario



Experiment Result: Statistic

	name	
Attack signature event name	SHELLCODE x86 inc ebx NOOP	1
	NETBIOS SMB-DS lsass unicode little endian overflow attempt	1
	NETBIOS SMB-DS lsass DsRoler UpgradeDownlevelServer unicode	1
	MISC MS Terminal server request	1
	NETBIOS DCERPC ISystemActivator path overflow attempt little endian	1
	DDOS Trin00 [illegible] default passwor	
	NETBIOS SMB [illegible] share access	
Victim IP	192.168.0.4	
	192.168.0.5	
	192.168.0.2	
	192.168.0.3	
Attacker IP	192.168.0.6	
	192.168.0.4	

Alert Correlation					
Total event	602				
TCP	11	ICMP	395	UDP	196
Correlated event	8	Reduction rate	98.70%		

associate ticket !! total number = 8

oid	source	target	class	signature name
1	192.168.0.6	192.168.0.4	probe	NETBIOS SMB-DS IPC$ unicode share access
2	192.168.0.6	192.168.0.4	Escalation	SHELLCODE x86 0x90 unicode NOOP
3	192.168.0.6	192.168.0.4	BreakIn	NETBIOS SMB-DS lsass DsRolerUpgradeDownlevelServer WriteAndX unicode little endian overflow attempt
4	192.168.0.6	192.168.0.4	BreakIn	NETBIOS SMB-DS lsass DsRolerUpgradeDownlevelServer unicode little endian overflow attempt
5	192.168.0.6	192.168.0.4	probe	MISC MS Terminal server request
8	192.168.0.4	192.168.0.3	DOS	DDOS Trin00 Master to Daemon default password attempt
7	192.168.0.4	192.168.0.2	DOS	DDOS Trin00 Master to Daemon default password attempt
6	192.168.0.4	192.168.0.5	BreakIn	NETBIOS DCERPC ISystemActivator path overflow attempt little endian unicode

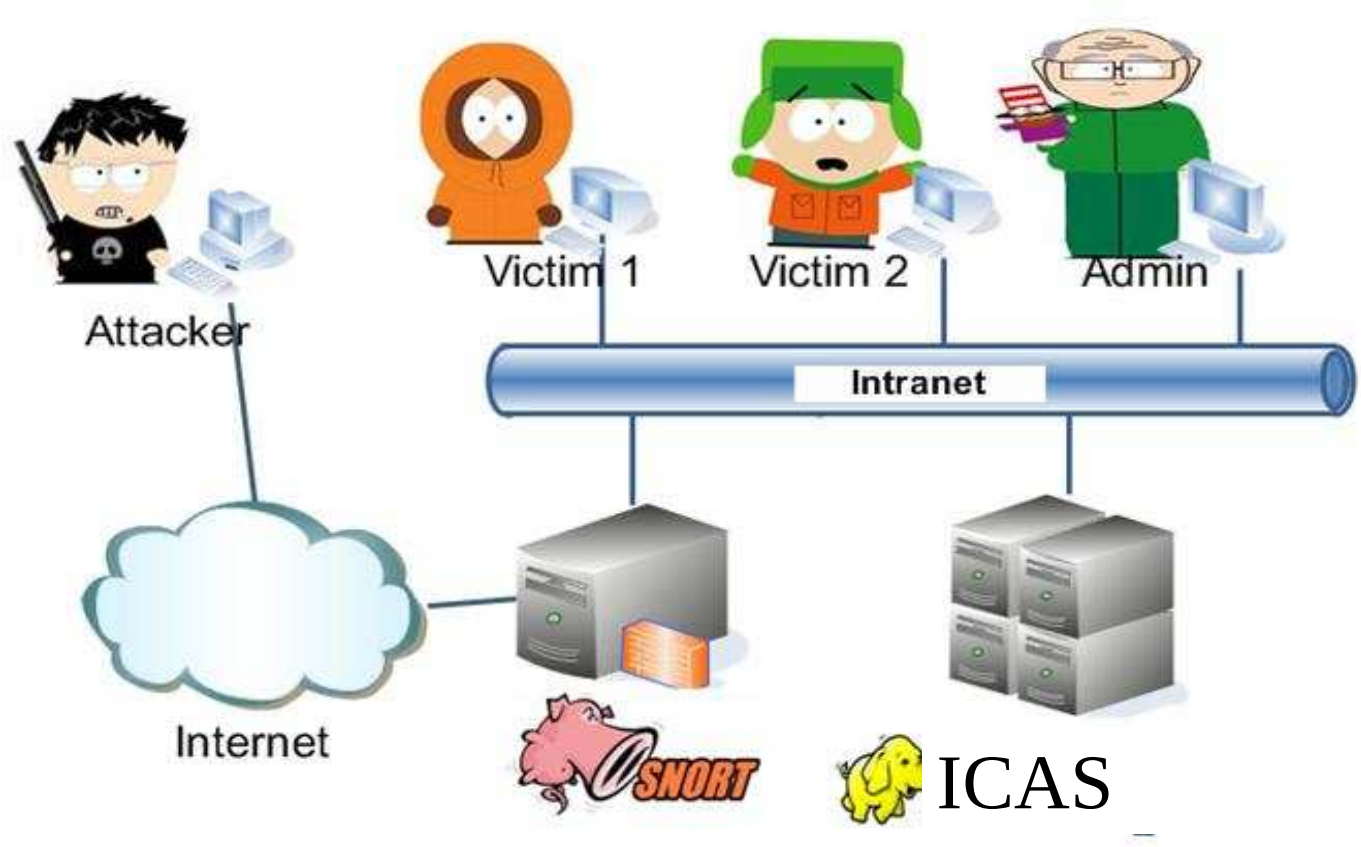
What's problem about the SEC ?

1. Enormous Data ⇔ less Efficient
2. Got Nothing if the database were crash
3. Memory and CPU Exhausted when system is running.

ICAS

- **ICAS, *IDS Cloud Analysis System***
- **Applying Cloud Computing technique**
  - Higher capability
  - Fault tolerance
- **Making alerts algorism to generate manifest report**
  - Reducing redundancy
  - Merge relation

ICAS Overview

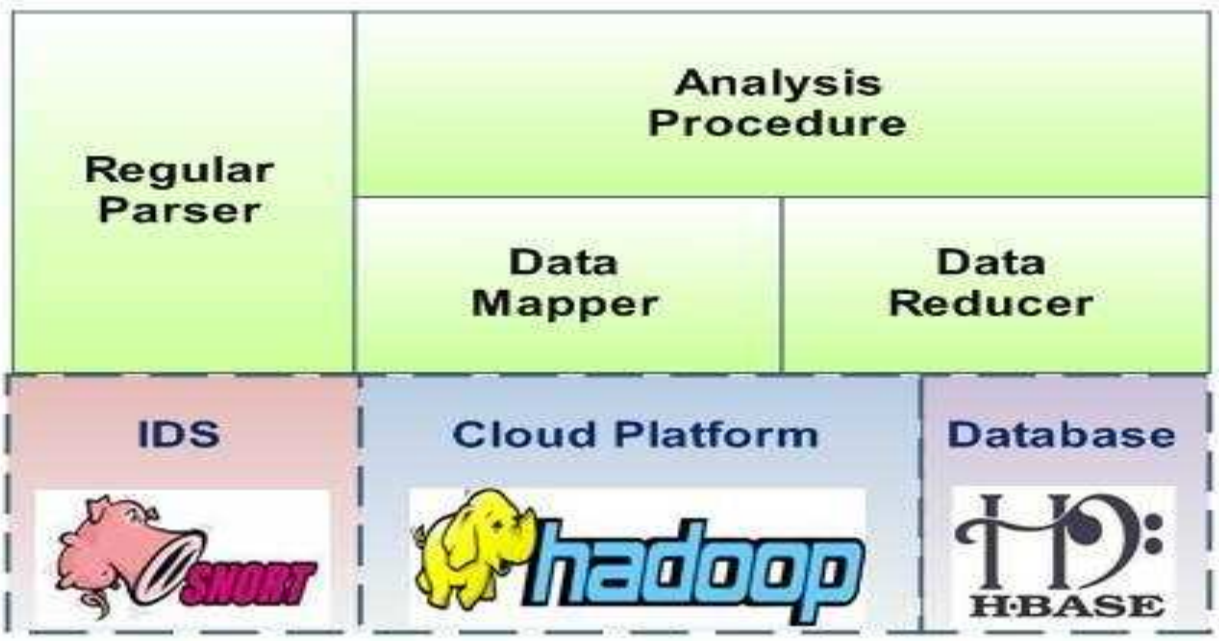


Change SEC to ICAS

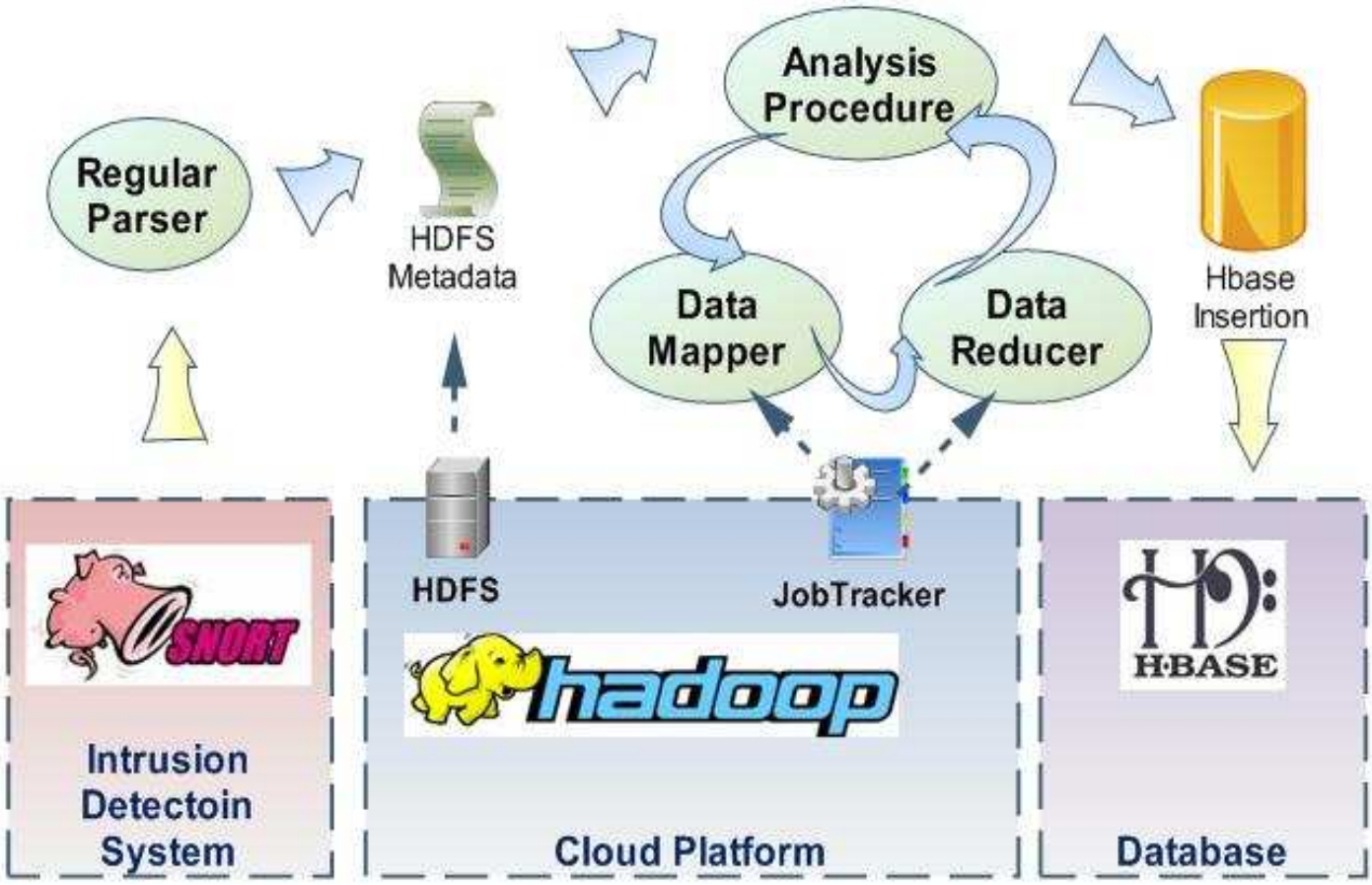
MySQL	→	HBase
Core Procedure	→	Map-Reduce
Linux	→	Hadoop + Linux
Single Machine	→	Multiple Machine

System Architecture

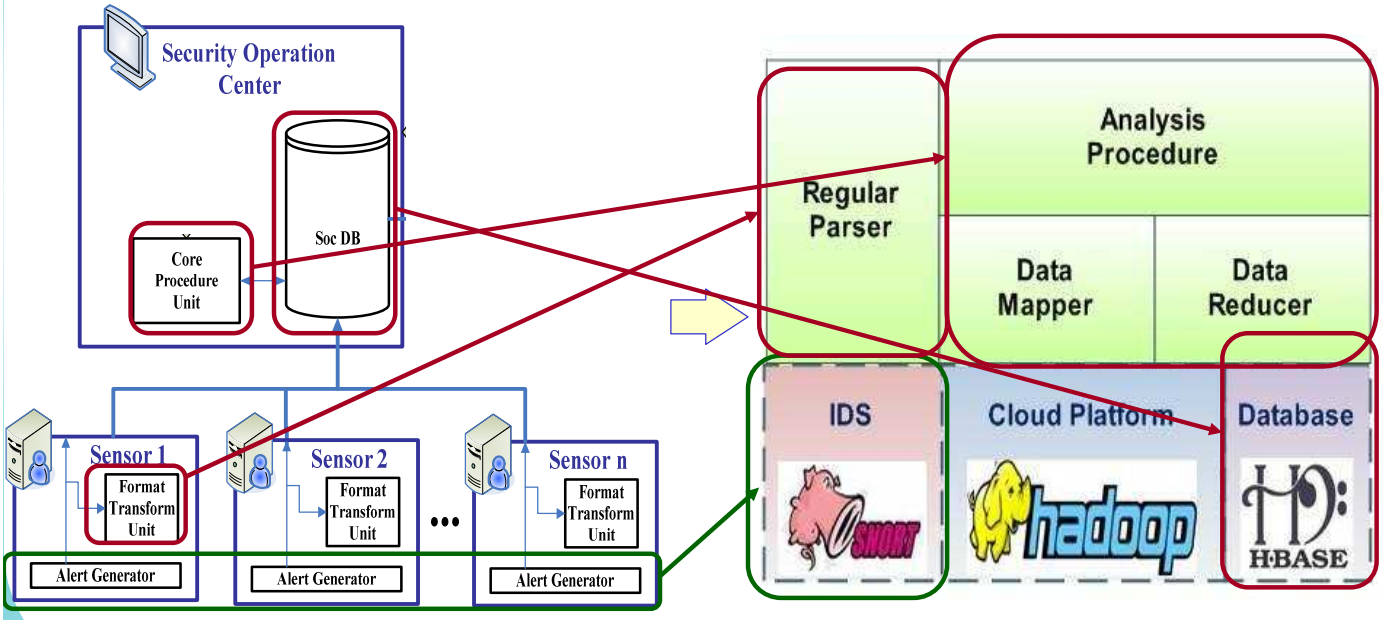
ICAS Component Overview

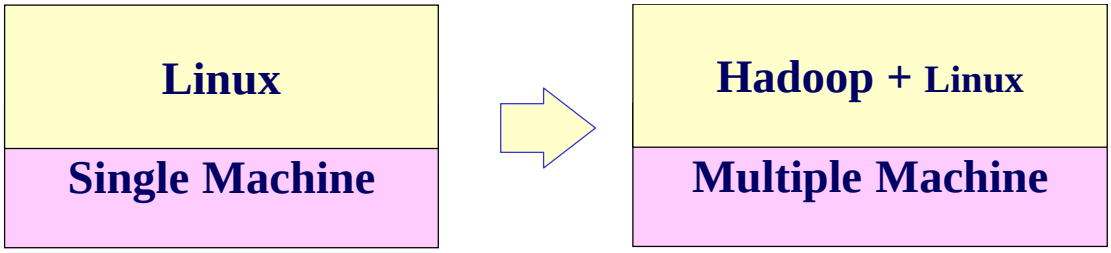


Program Procedure



Change SEC to ICAS





- **Applications**
  - MySQL -> HBase
- **Program language**
  - PHP -> JAVA & JSP
- **Data transfer**
  - Snort to MySQL -> Log to HDFS

Core Procedure

- **Format Transfer Unit**
  - Setup Snort logging to MySQL
  - Setup MySQL client logging to remote MySQL server
- **Core Procedure Unit**
  - Fuse redundant data
  - Merge data as event

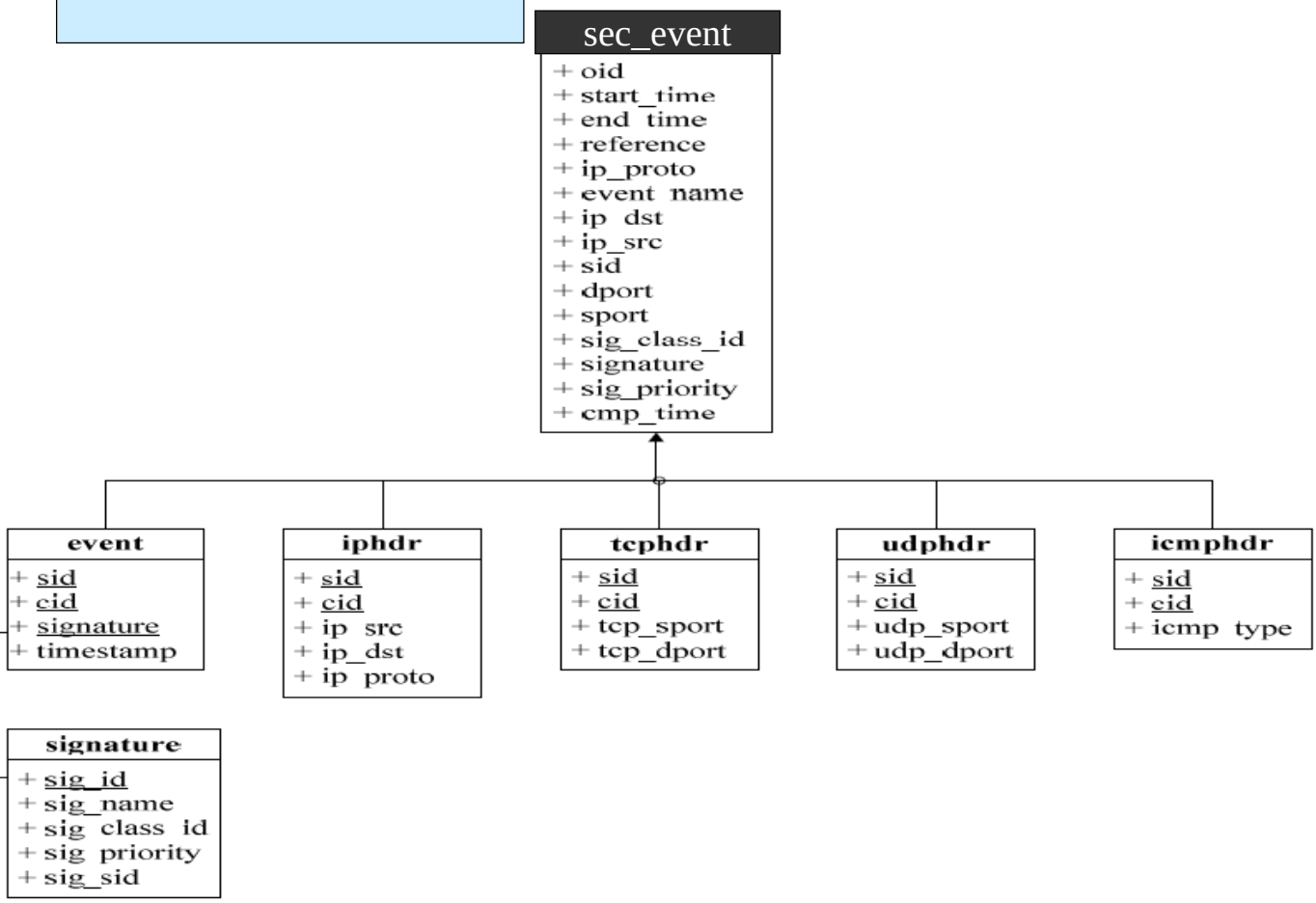
Map-Reduce

- **Regular Parser**
  - Parsing original snort log and transfer to HDFS (hadoop file system)
- **Analysis Procedure**
  - Dispatch job if pool is not empty and insert the result into database
- **Data Mapper**
  - <key, value> mapping
- **Data Reducer**
  - <"key1", value1…valueN>
  - <"key2", value1…valueN>

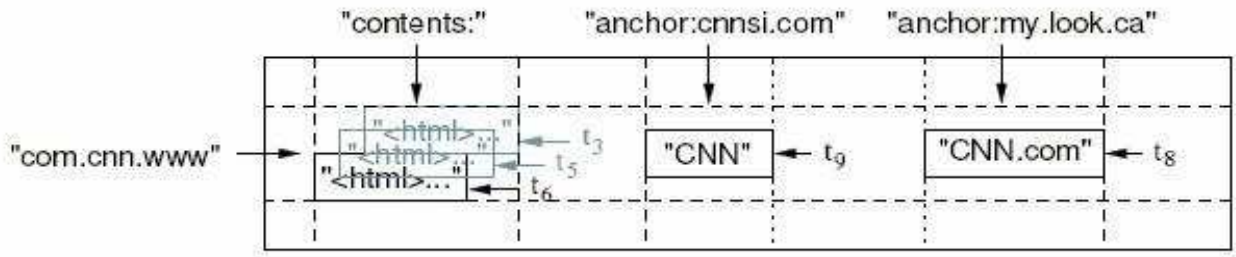
MySQL

Tables name	Correlated events		
description	The final integrated events		
Primary key	oid		
Field	**Type**	**NULL**	**comment**
oid	Int(10)	N	Object id
Start_time	Datetime	N	The start time of this incident
End_time	Datetime	N	The end time of this incident
Reference	Varchar(255)	Y	Recode the merged alerts key
IP_proto	Varchar(255)	Y	IP protocol
Event_name	Varchar(255)	Y	The signature event name
IP_dst	Varchar(255)	Y	Destination IP (only one)
IP_src	Varchar(255)	Y	Source IP (various)
Sid	Varchar(255)	Y	Sensor id
Dport	Varchar(255)	Y	Destination port
Sport	Varchar(255)	Y	Source port
Sig_class_id	Tinyint(3)	Y	Our proposed signature class
Signature	Varchar(255)	Y	Snort signature id number
Sig_priority	Tinyint(3)	Y	Signature priority

MySQL



HBase



Row Key	Time Stamp	Column *"contents:"*	Column *"anchor:"*		Column *"mime:"*
com.cnn.www	**t5**		**"anchor:cnnsi.com"**	**CNN**	
	t4		**"anchor:my.look.ca"**	**CNN.com**	
	t3	**<html>...**			**text/html**
	t2	**<html>...**			
	t1	**<html>...**			

HBase

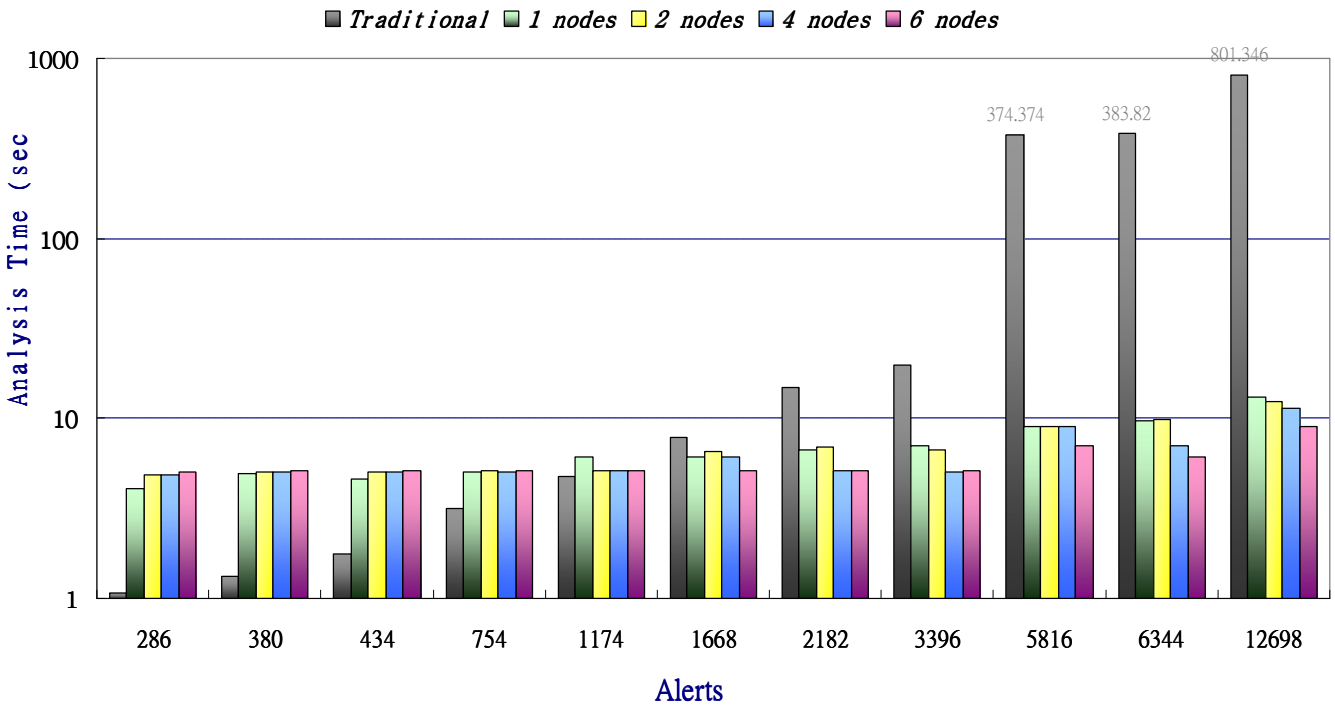
Row Key	Time Stamp	Column "signature:"	Column "Infor:"		Column " SourceIP :"
(Destination IP) dIP1	t1	sig1	Infor :Port	p1,p2..	sIP1
	t2	sig2	Infor : ... (the other info)	o1,o2…	sIP2
dIP2	…	…	Infor :…	..	…
…	…	…	…	..	…

Experiment

- **Machine:**
  - CPU : Intel quad-core, Memory : 2g,
- **OS :** Linux : Ubuntu 8.04 server
- **Software : version**
  - Hadoop : 0.16.4
  - Hbase : 0.1.3
  - Java : 6
- **Alerts Data Sets**
  - MIT Lincoln Laboratory, Lincoln Lab Data Sets
  - Computer Security group at UCDavis, tcpdump file

Experimental Result

The Consuming Time of Each Number of Data Sets



Experimental Result

Throughput Data Overall

Origianl Alerts	Analysis Time (sec)					Results	Reduction Rate
	Traditional	1 nodes	2 nodes	4 nodes	6 nodes		
286	1.068	4.087	4.869	4.864	5.077	30	89.51%
380	1.333	4.94	5.069	5.067	5.097	11	97.11%
434	1.76	4.61	5.066	5.068	5.09	9	97.93%
754	3.145	5.066	5.079	5.038	5.096	16	97.88%
1174	4.73	6.066	5.093	5.089	5.097	33	97.19%
1668	7.909	6.07	6.56	6.071	5.082	16	99.04%
2182	14.949	6.671	6.95	5.166	5.088	16	99.27%
3396	19.901	7.053	6.654	5.076	5.091	68	98.00%
5816	374.374	9.081	9.076	9.07	7.076	66	98.87%
6344	383.82	9.68	9.872	7.069	6.069	72	98.87%
12698	801.346	13.096	12.367	11.367	9.083	36	99.72%

Pros & Cons

- Efficient
- Scalable
- Economical
- Reliable
- Non-realtime
- Latency
- Immature

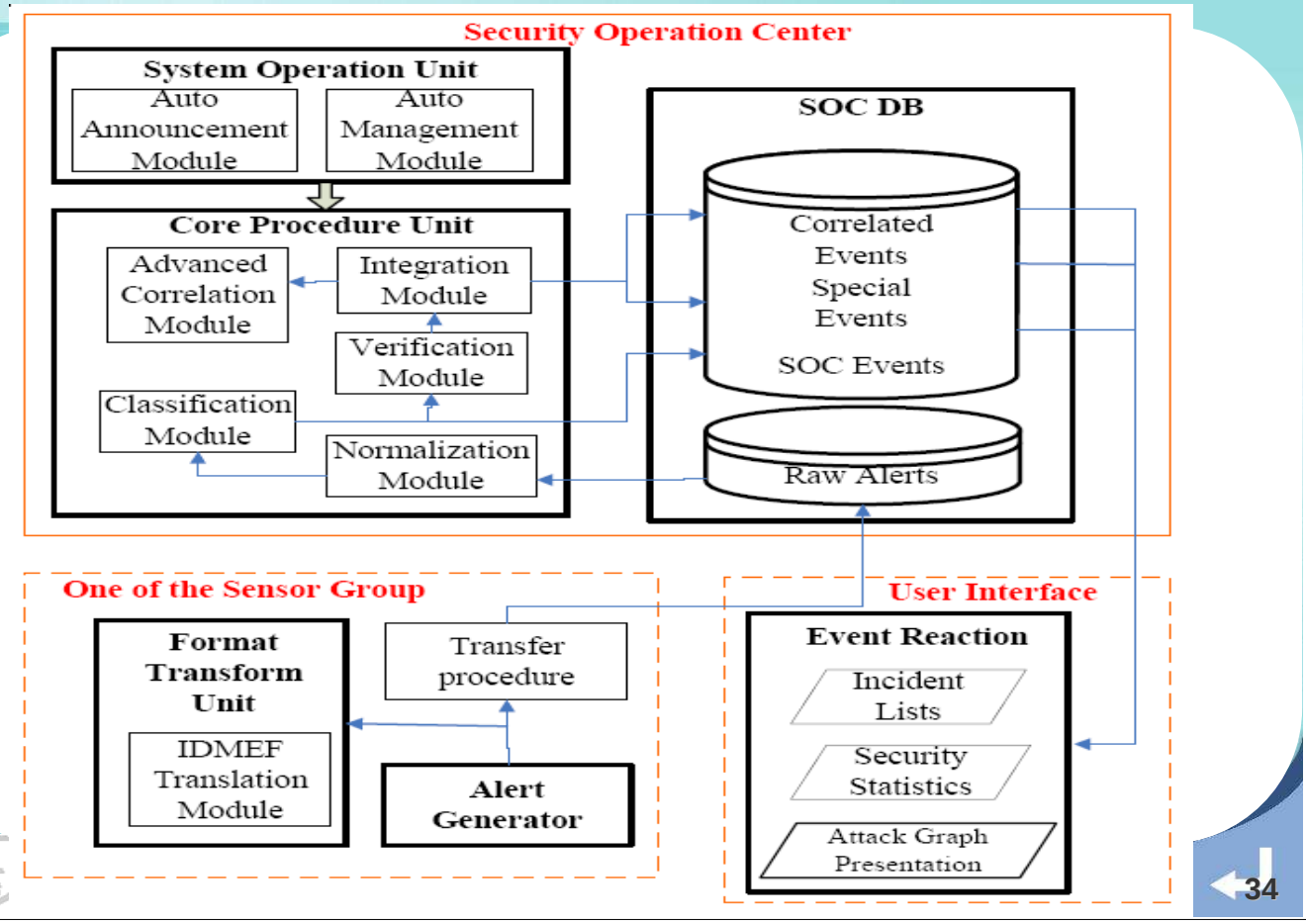
Hadoop Development Issues

- **Fully based on correct Hadoop's Version (Neither backward nor upward compatibility)**
  - ICAS can work on Hadoop 0.16.4.
  - ICAS has 8 errors and 8 deprecations on Hadoop 0.18.3
  - ICAS has 26 errors and 22 deprecations on hadoop 0.20.0
  - A word-count sample code on hadoop 0.20 can't work for hadoop 0.18
  - HBase's "A" version is only correspond to Hadoop's "A" version (upper or lower not work)
- **Sample codes are hardly to find**
- **Deeply in Object-oriented programming**

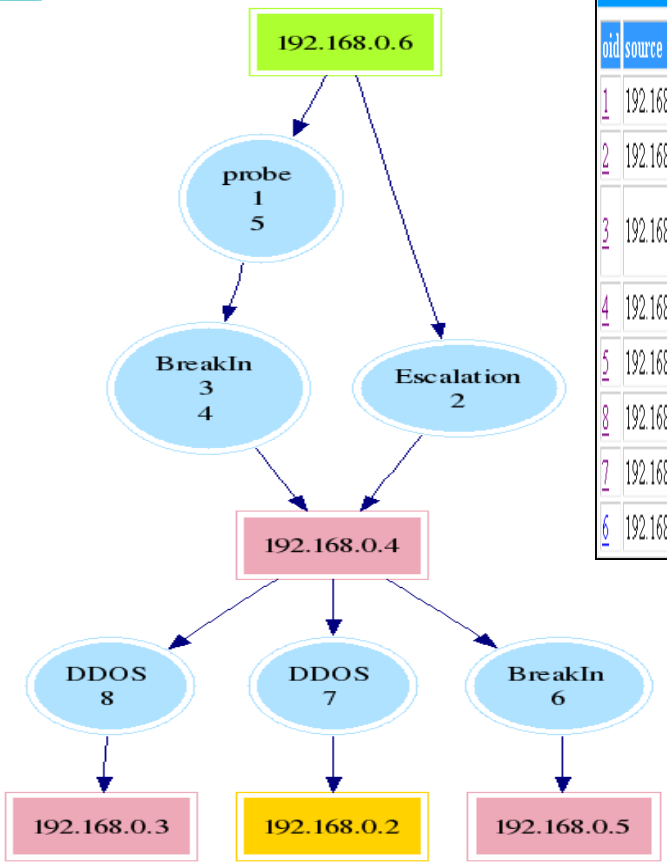
Conclusions

- ICAS supplies a efficient way to analyze and merge huge number of alerts based on cloud platform
- Until now, there are many components needed to implement

Future Works : Overview



Future Work : Final Result



associate ticket !! total number = 8

bid	source	target	class	signature name
1	192.168.0.6	192.168.0.4	probe	NETBIOS SMB-DS IPC$ unicode share access
2	192.168.0.6	192.168.0.4	Escalation	SHELLCODE x86 0x90 unicode NOOP
3	192.168.0.6	192.168.0.4	BreakIn	NETBIOS SMB-DS lsass DsRolerUpgradeDownlevelServer WriteAndX unicode little endian overflow attempt
4	192.168.0.6	192.168.0.4	BreakIn	NETBIOS SMB-DS lsass DsRolerUpgradeDownlevelServer unicode little endian overflow attempt
5	192.168.0.6	192.168.0.4	probe	MISC MS Terminal server request
8	192.168.0.4	192.168.0.3	DOS	DDOS Trin00 Master to Daemon default password attempt
7	192.168.0.4	192.168.0.2	DOS	DDOS Trin00 Master to Daemon default password attempt
6	192.168.0.4	192.168.0.5	BreakIn	NETBIOS DCERPC ISystemActivator path overflow attempt little endian unicode

Future Works

- Including more IDS logs
- Graphical final result
- Prepare more large-scale and complete experiment

Thank You !

&

Question ?