



# Achieving Real Real-Time Context-Based Actionable Intelligence in Cyber Investigations

**Joel Ebrahimi**  
**Senior Solutions Architect**  
**Bivio Networks Inc**

# LEA Investigations in Cyberspace

---

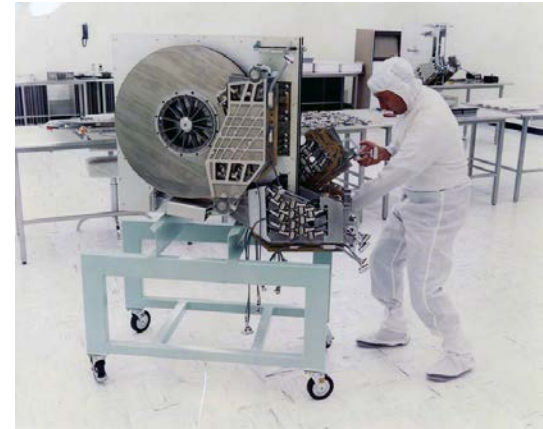
- /// Cyberspace now the dominant medium of communication
- /// Threats and breaches growing yet cyber forensic and investigative technologies lagging behind
- /// Cyber Investigations present significant challenges:
  - Massive amount of bandwidth
  - New and constantly evolving applications
  - New devices
  - New networks



# Today's Tools

---

- /// Full Packet Capture
- /// Server Logs
- /// Meta data extraction
  - Stove Pipe Collection
- /// Analyzing Platforms



# Data Retention System

---

- /// Cyber equivalent of CCTV cameras
- /// Establish complete record of networking domain
  - Captures every communication for all users
- /// Sample use cases
  - Identify individuals engaged in criminal behaviour
  - Investigate insider or outsider threats to corporate assets
  - Investigate relationships between individuals using cyber space
  - Maintain record of network traffic for compliance, intelligence or auditing



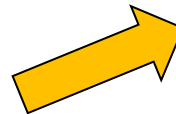
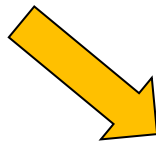
Provide information to answer simple question:  
***“Who did what, when, where and with whom?”***


# Mass Metadata Extraction



```
6000 0000 034d 063c 2001 4860 b002 0000
0000 0000 0000 0068 2001 0470 1f03 0233
0000 0000 0000 0002 0050 2433 1a11 7590
88bd 4d54 5019 1920 8732 0000 653c 2f61
3e3c 2f66 6f6e 743e 3c70 3e3c 666f 6e74
2073 697a 653d 2d32 3e26 636f
```

```
6000 0000 034d 063c 2001 4860 b002 0000
0000 0000 0000 0068 2001 0470 1f03 0233
0000 0000 0000 0002 0050 2433 1a11 7590
88bd 4d54 5019 1920 8732 0000 653c 2f61
3e3c 2f66 6f6e 743e 3c70 3e3c 666f 6e74
2073 697a 653d 2d32 3e26 636f
```

```
6000 0000 034d 063c 2001 4860 b002 0000
0000 0000 0000 0068 2001 0470 1f03 0233
0000 0000 0000 0002 0050 2433 1a11 7590
88bd 4d54 5019 1920 8732 0000 653c 2f61
3e3c 2f66 6f6e 743e 3c70 3e3c 666f 6e74
2073 697a 653d 2d32 3e26 636f
```



Protocol	Twitter	
Application	Twitter	
UserID	Protocol	HTTP
Password	Application	Facebook
Following	UserID	Joeevil
Members	Password	Protocol
Posts	Attributes	Application
...	Friends	UserID
...	...	Password
...	...	Chatroom
...	...	Members
...	...	Chat size
...	...	...
...	...	...

*100s of protocols with 1000s of metadata attributes  
turn network activity into powerful searchable medium*

# Correlation: Finding the Needle in a Haystack

---

- /// Mapping electronic to physical world harder than ever
- /// Data resides in multiple domains
- /// Creating actionable information often relies on synthesis of multiple pieces of data



# Correlation Example

## Signaling Flows

L7 Flow Records

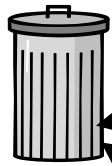
Field	Field	Protocol
John_doe	124.56.31.212	RADIUS ADSL
124.56.31.212	00:11:22:33	RADIUS WiMAX
124.56.31.212	IMEI	RADIUS 3G

HTTP related Metadata

Src IP Address	Dst IP Address	Protocol	URL
124.56.31.212	1.2.3.4	HTTP	www.blog.com

VoIP related Metadata

SRC Phone	DST Phone	Protocol	Src IP
+4411223344	+97122334455	SIP	124.56.31.212



Merging Phase

Merging Phase

Useless fields are filtered & deleted


John_Doe	1.2.3.4	HTTP	www.blog.com	124.56.31.212
00:11:22:33	+4411223344	IMEI	+97122334455	124.56.31.212

**Result:** → **Unified Correlated Records**

# Packet Capture Integration

- Enhances Data Retention records with full payload option
- Provides for analysis with 3<sup>rd</sup> party tools
- Enables full session reconstruction and playback
- Massively scalable: data amount captured limited only by economics
- First system to make full packet capture viable in large scale

Black_Cat	1.2.3.4	HTTP	www.blog.com	124.56.31.212	PAYLOAD
00:11:22:33	+4411223344	IMEI	+97122334455	124.56.31.212	



```
00000000 61 6e 74 61 67 6f 6e 69 73 74 0a 64 0a 34 0a 30 antagoni st.d.4.0
00000010 57 51 67 64 0a 31 34 31 0a 1c a2 04 08 ce 39 6e WQgd.141 .....9n
00000020 d2 51 9b 0f de cb 6f ac e8 7c 02 08 06 05 47 2c .Q....o...|...G,
00000030 15 60 16 e4 a6 a7 13 61 56 8d b0 f8 29 8e 18 45 .....a V...)..E
00000040 38 11 2a 5a 68 7d 3c c2 df 13 54 8a 06 d4 a5 60 8.*Zh}<...T....
00000050 7c 46 5b 70 1d 78 1e 23 6f 8d f5 7c 2d 85 dc 17 |F[p.x.# o..|-...
00000060 6f 58 3c f0 62 d2 d8 1a 6b 9f 03 fc e8 b4 95 74 oX<.b... k.....t
00000070 4c c6 97 2d 05 15 26 db 4a 8b f8 9c 8c ce 78 1f L....&. J....x.
00000080 cb 17 4a 18 f6 54 c1 28 a5 9c 1b ab 9d 28 80 fb ...J..T.( .....(
00000090 5c 72 e4 b2 a8 dd 79 fa 47 00 01 55 0a 16 07 07 \r....y. G..U....
000000A0 09 31 37 24 68 6e 62 0a 31 30 0a .....17$hnb. 10.
```



# The Result: "LEA Search Engine"

- Real-time response
- Simple but effective GUI for queries
- Actionable information

The screenshot displays the LEA Search Engine interface. The main window shows a query titled "raw data wimax\_user between 2011/07/01 00:00:00 and 2011/07/08 09:15:00". The results are displayed in a table with columns: nas\_ip, sub\_id, user\_ip, user\_netmask, user\_mac, cpe\_mac, and bsid. A "Show Info" popup window is open over the first row, displaying user details: subscriber: ua278r, name: Vadk, surname: Hrika, account: flat, birth\_place: Sardnerie, and birth\_date: 1960-6-4. The "Available Queries" list on the left includes "raw data wimax\_user".

nas_ip	sub_id	user_ip	user_netmask	user_mac	cpe_mac	bsid
77.85.255.252	ua278r	77.85.122.82	255.255.255.255	00519f96e019	005165d092fd	00062ca4ff
77.85.255.252	u028r	77.85.246.246	255.255.255.255	0051b0fca914	0051340efa83	00062ca4ff
77.85.255.252	ua0h3y	77.85.207.70	255.255.255.255	00517a7abc60	00514998baea	00062ca4ff
77.85.255.252	u50fn3R	77.85.120.156	255.255.255.255	00512a904872	0051e008ba68	00062ca4ff
77.85.255.252	u0Fn0y	77.85.207.126	255.255.255.255	0051ef0bc5ae	0051edf5a23b	00062ca4ff
77.85.255.252	u0Dm0y	77.85.253.79	255.255.255.255	00513f20297f	0051d517f4e1	00062ca4ff
77.85.255.252	u0D73R	77.85.232.66	255.255.255.255	005160e1e758	0051dbee50e0	00062ca4ff
77.85.255.252	u0D735	77.85.207.113	255.255.255.255	00512a0b53ae	0051bc48b669	00062ca4ff
77.85.255.252	u52m0j	77.85.197.22	255.255.255.255	0051146a728e	0051f9a3a620	00062ca4ff
77.85.255.252	u52a35	77.85.152.34	255.255.255.255			
77.85.255.252	u5Dra5	77.85.213.159	255.255.255.255			

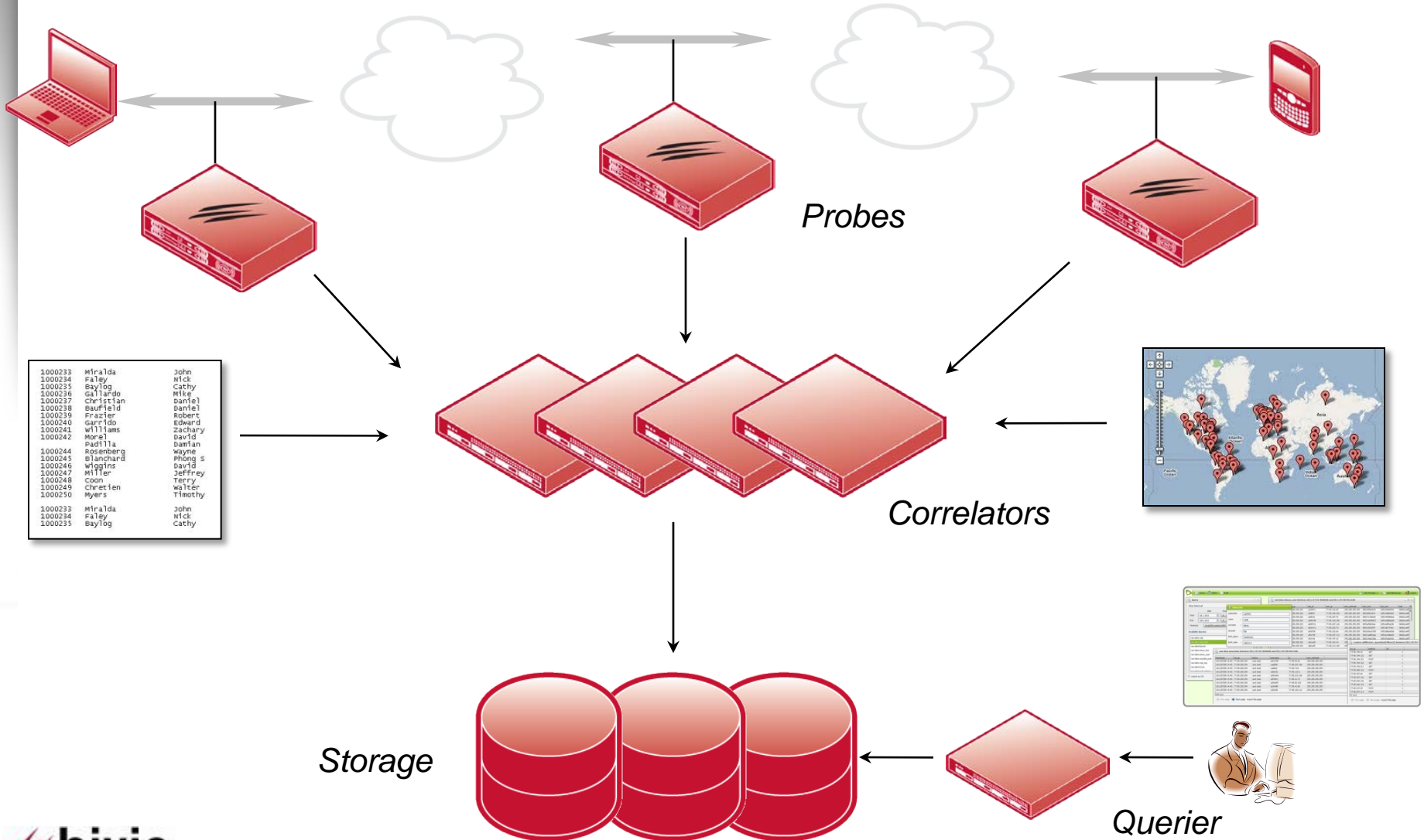
Below the main table, there is another query titled "raw data associator between 2011/07/01 00:00:00 and 2011/07/08 09:15:00". Its results table has columns: timestamp, nas\_ip, status, username, ip, and user\_netmask.

timestamp	nas_ip	status	username	ip	user_netmask
2011/07/05 14:45:	77.85.255.254	acct start	u02779R	77.85.95.26	255.255.255.255
2011/07/05 14:45:	77.85.255.254	acct start	ua883R	77.85.207.106	255.255.255.255
2011/07/05 14:45:	77.85.255.254	acct start	ua0a0j	77.85.7.63	255.255.255.255
2011/07/05 14:45:	77.85.255.254	acct start	u50m0y	77.85.123.6	255.255.255.255
2011/07/05 14:45:	77.85.255.254	acct start	u0D4e9y	77.85.215.186	255.255.255.255
2011/07/05 14:45:	77.85.255.254	acct start	u0D48y	77.85.22.72	255.255.255.255
2011/07/05 14:45:	77.85.255.254	acct start	u0D4af5	77.85.50.169	255.255.255.255
2011/07/05 14:45:	77.85.255.254	acct start	u0D4a9R	77.85.42.86	255.255.255.255
2011/07/05 14:45:	77.85.255.254	acct start	u52h3R	77.85.183.112	255.255.255.255

On the right side, there is a "counter" window showing a table with columns: src\_ip, method, and all.

src_ip	method	all
77.85.194.32	GET	1
77.85.194.122	GET	2
77.85.194.151	POST	1
77.85.194.162	GET	1
77.85.194.211	GET	1
77.85.205.143	POST	1
77.85.207.66	GET	1
77.85.207.126	GET	1
77.85.245.132	GET	1
77.85.246.115	GET	1
77.85.247.20	POST	1
77.85.247.113	POST	1

# DRS Architecture



# Technology Highlights

---

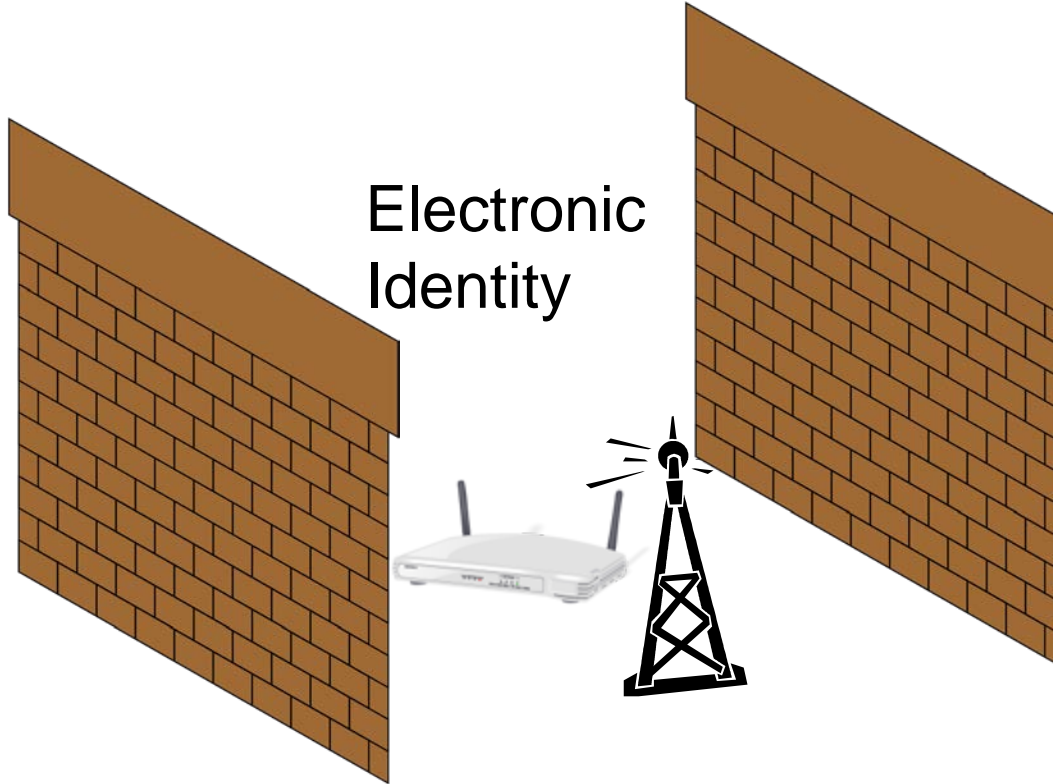
- /// Mass metadata extraction for contextual information
- /// Multi-source data collection and correlation
- /// Instantaneous query response
- /// Wired and mobile networking support
- /// Scalability to national-level deployments
- /// Ease of integration with legacy and 3<sup>rd</sup> party systems
- /// Packet capture options for full payload collection

# LEA Challenge: Physical Identification

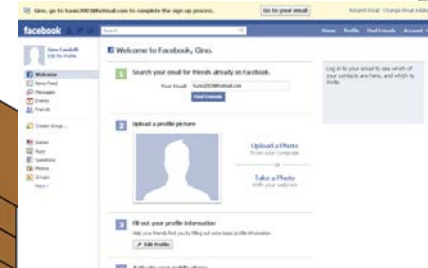
Physical Identity



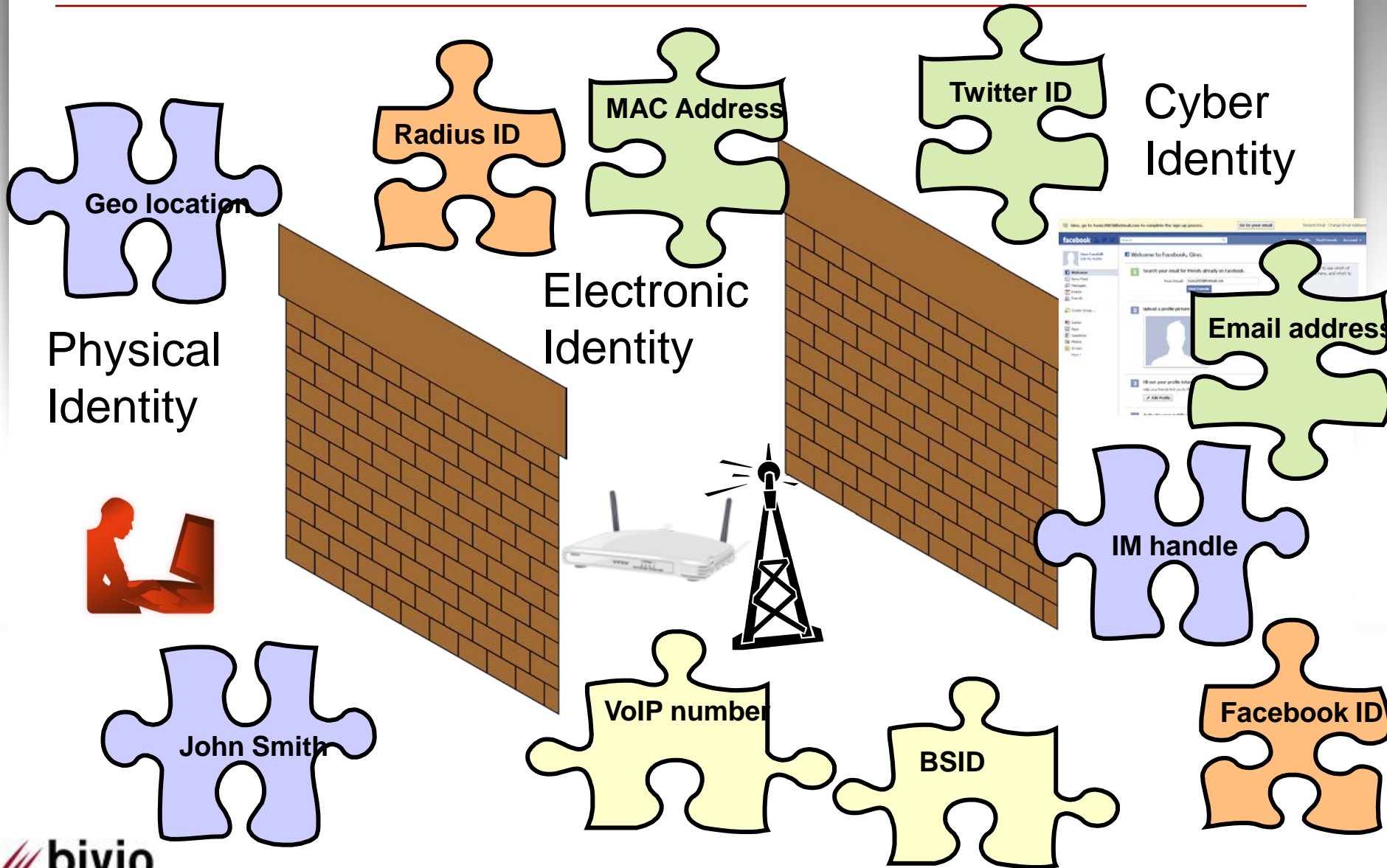
Electronic Identity



Cyber Identity



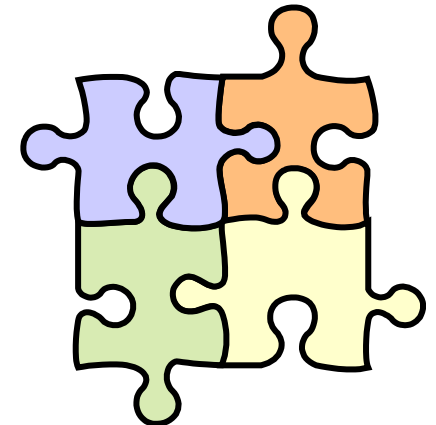
# Different pieces of the same puzzle!



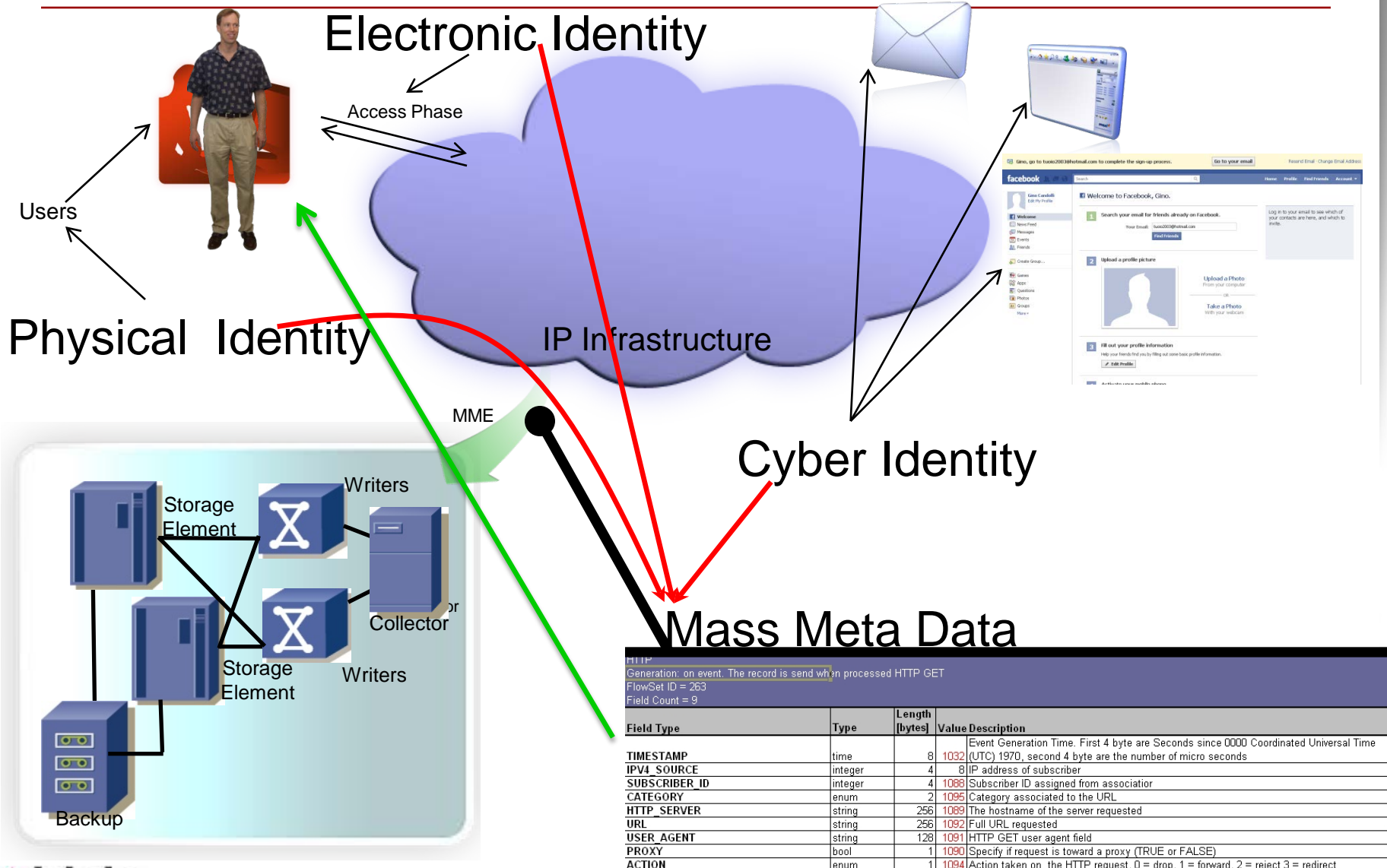
# DRS Correlation: Solving the Puzzle

---

- /// Mixes and merges information coming from different sources (probes, network elements, static DB)
- /// Dynamically creates the links between the various entities (Cyber Identity, Electronic Identity and Physical Identity)
- /// Enables real-time synthesis of actionable information
- /// **Converts fragmented data into meaningful, actionable intelligence: who, what, where and when**



# Finding the Physical Identity



HTTP

Generation: on event. The record is send when processed HTTP GET

FlowSet ID = 263

Field Count = 9

Field Type	Type	Length [bytes]	Value	Description
TIMESTAMP	time	8	1032	Event Generation Time. First 4 byte are Seconds since 0000 Coordinated Universal Time (UTC) 1970, second 4 byte are the number of micro seconds
IPV4_SOURCE	integer	4	8	IP address of subscriber
SUBSCRIBER_ID	integer	4	1088	Subscriber ID assigned from associator
CATEGORY	enum	2	1095	Category associated to the URL
HTTP_SERVER	string	256	1089	The hostname of the server requested
URL	string	256	1092	Full URL requested
USER_AGENT	string	128	1091	HTTP GET user agent field
PROXY	bool	1	1090	Specify if request is toward a proxy (TRUE or FALSE)
ACTION	enum	1	1094	Action taken on the HTTP request, 0 = drop, 1 = forward, 2 = reject 3 = redirect

# Summary

---

- /// The key enabler for real time Cyber Investigations
- /// Captures, synthesizes and retains meaningful intelligence for every communication across the network
- /// Provides the evidence needed to look back in time and identify who did what, when, where and with whom
- /// Designed for the rapidly evolving Web 2.0 world
- /// Real-time Intelligence
- /// For any network size and any retention time window

