

Fighting Identity Theft

Big Data Analytics to the Rescue

Seshika Fernando

WSO2

Me - Seshika

- Computer Science & Finance
- Streaming Analytics



- 100% Open Source Middleware Company
- Apache Way
- <http://wso2.com/>

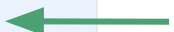




Quantified

- **\$2.5m** per Enterprise
- **#1** Consumer Complaint
- Every **2** seconds
- **51%** Enterprises use Big Data Analytics

Service Provider



Identity Providers

User



Log in with Facebook
Log in with Google
Log in with GitHub
Log in with OpenStack

Or

[LOG IN](#) » [RETURN TO EVENTS](#)

I already have a Linux Foundation ID
 I need to create a Linux Foundation ID

seshika

.....|


Enter the password that accompanies your username.

Log in


[Request new password](#)

Favourites


Web Apps




Client Store
1.0.0
★★★★★




Concur Financial System
1.0.0
★★★★★




Dinner on Demand
1.0.0
★★★★★




Jenkins Build Server
1.0.0
★★★★★




Marketing Dashboard
1.0.0
★★★★★




Open Stack
1.0.0
★★★★★




PeopleHR
1.0.0
★★★★★




Product Performance Te...
1.0.0
★★★★★




Redmine
1.0.0
★★★★★



Support Portal
1.0.0
★★★★★



WSO2 Salesforce
1.0.0
★★★★★



WSO2 Travel App
1.0.0
★★★★★

Authentication Analytics

- Blacklisted IP address
- Single IP, multiple users
- Single user, multiple IPs
- Login from new IP address
- Abnormal frequency of logins
- Abnormal login times
- Multiple login failures
- Multifactor authentication failures



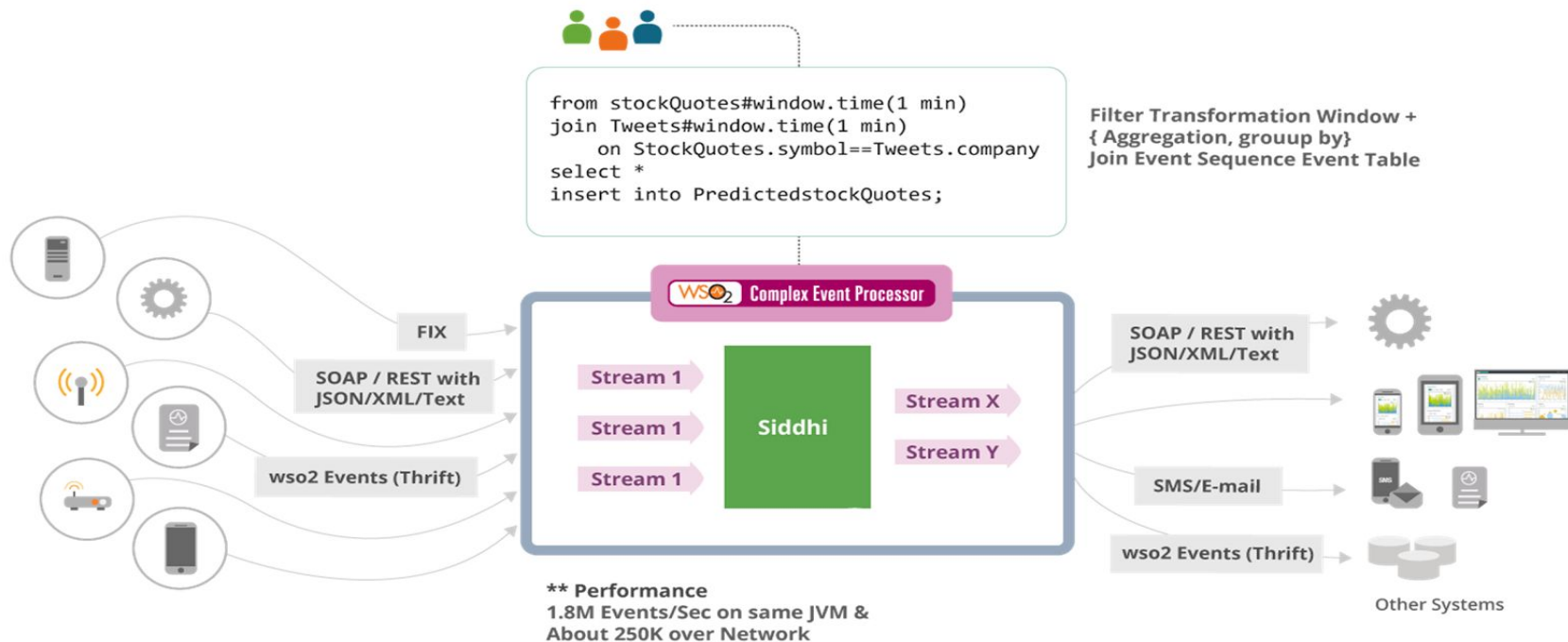
“Should I arrest Clark Kent for identity theft or should Clark Kent have me arrested for identity theft? This is all so very confusing!”

Authorization Analytics

- User/Role accessing a new resource
- Abnormal resource access frequency
- Access denied for multiple resources, for the same user
- Abnormal usage frequency of high privilege accounts
- High risk privilege escalation



Complex Event Processing



** Notify if there is a 10% increase in overall trading activity AND the average price of commodities has fallen 2% in the last 4 hours*

Blacklists

```
define table BlacklistedIPTable (ipAddress string);

from loginStream[ (ip == BlacklistedIPTable.ip) in BlacklistedIPTable ]
select *
insert into alertStream;
```

Whitelists

```
define table IPTable (ipAddress string);

from loginStream[ not(ip == IPTable.ip) in IPTable ]
select *
insert into alertStream;
```

Counting

```
from loginFailureStream#window.time(1 hour)
select username, count(timestamp) as loginFailCount
group by username
having loginFailCount > 30
insert into alertStream;
```

1 to many relationships

```
from      e1 = loginStream ->
          e2 = loginStream[(e1.ip == e2.ip) and (e1.username != e2.username)] <2:>
          within 1 day
select e1.ip, e1.username, e2[0].username, e2[1].username
insert into alertStream;
```

Adaptive Analytics

User Profiling (UEBA)

- Time
- IP/Geo-location
- Frequency
- Typing Patterns
- Service Provider(s)
- Identity Provider(s)

Condescending Wonka

Stats

Birthday: June 30, 1971

Join Date: January 13, 2015

Last Activity: February 24, 2015

Total Entries: This user has not published any entries yet.

Total Comments: This user has not posted any comments yet.

Total Reviews: 20 [view all](#) →

Total Favorites: 15 (9 entries, 6 members) [view all](#) → [More on Favorites functionality](#) →

Profile

Website: <http://knowyourmeme.com/memes/condescending-wonka-creepy-wonka>

Location: Unnamed European town

Occupation: Candy-maker

Interests: The Everlasting Gobstopper

Bio:
Condescending Wonka is an advice animals image macro series featuring a screen capture of actor Gene Wilder in the 1971 musical Willy Wonka and the Chocolate Factory. As its name suggests, the captions can be characterized as patronizing and condescending.

Wonka usually logs in between **8am - 10am**, from an IP address in **Chicago**, and logs into **Redmine** and **Concur**, using his **Google** Credentials

Behavioural Rules

- Based on
 - Time
 - Login Frequency
 - Geo Location
 - List of Service Providers
 - List of IDPs

```
from loginStream#window.time(1 hour) as str join loginCountTable as tbl
on str.username == tbl.username
select str.username, count(str.timestamp) as curLoginCount, tbl.maxLoginCount
group by str.username
having curLoginCount > maxLoginCount
insert into alertStream;
```

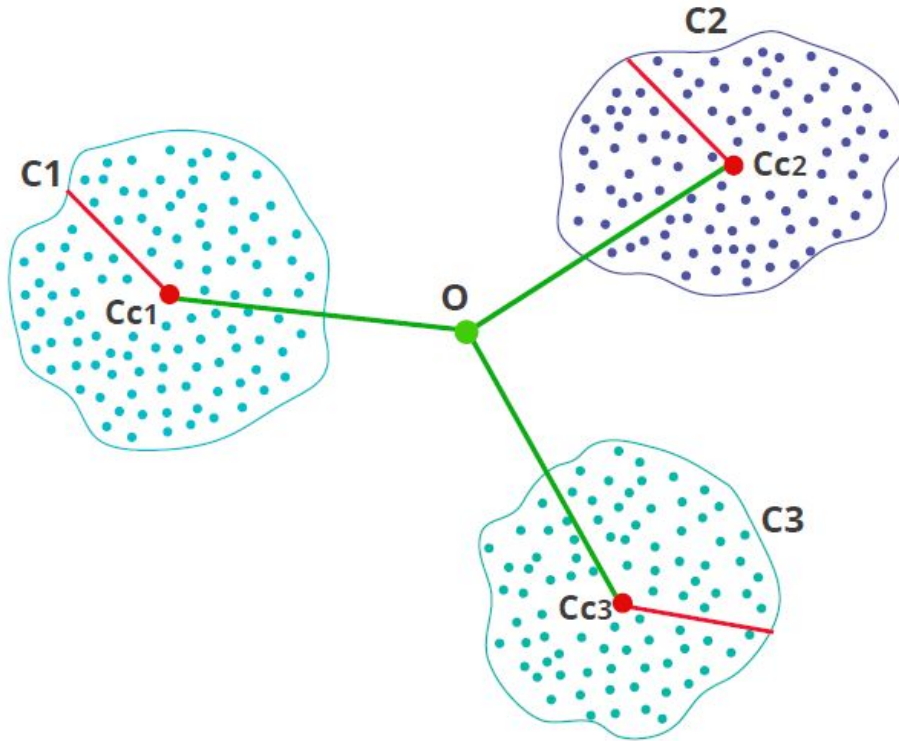
Scoring

- Use combination of rules
- Give weights to each rule
- Single number to represent suspicion through multiple indicators
- Use a threshold to identify anomalies

$$\text{Score} = w1 * \text{time} + w2 * \text{frequency} + w3 * \text{location} + w4 * \text{SPs} + w5 * \text{IDPs}$$



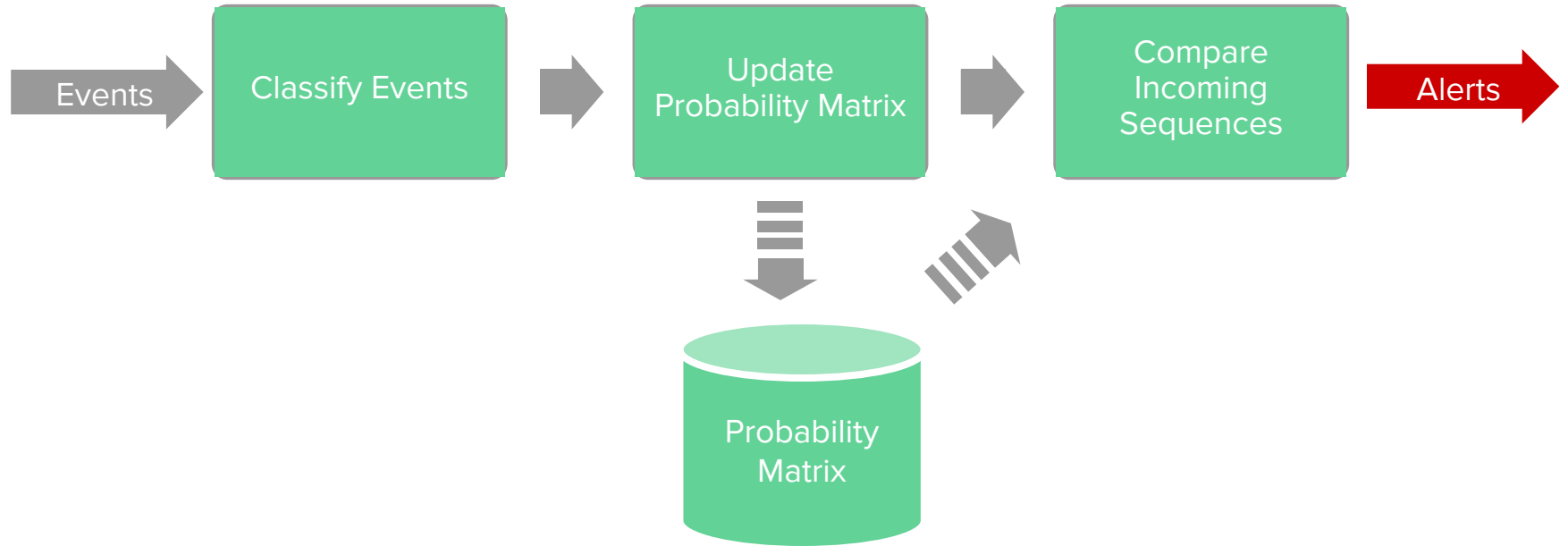
Clustering



Features

- Time
- Geo Location
- IdP
- SP Type

Markov Models



Audit Trail Analytics

The screenshot displays the 'Audit Comparator' application window. At the top, it shows the time range 'From 8:00:00' to 'To 12:00:00' on '10/10/2016'. Below this, there are filters for 'Service Provider' (PeopleHR), 'User' (Paul), and a 'Select' dropdown. The main area is divided into three columns. The left column contains a vertical timeline with three blue arrows pointing downwards. The middle column shows a vertical timeline with two orange arrows pointing downwards, each associated with a yellow callout box. The first callout box contains the text: 'SP Permission Added', 'User: Admin', 'Date: 10/10/2016', and 'Time: 9:50 am'. The second callout box contains the text: 'SP Permission Revoked', 'User: Admin', 'Date: 10/10/2016', and 'Time: 10:10 am'. The right column is currently empty. On the left side of the application, there is a vertical toolbar with icons for user profile, settings, folder, search, print, and delete.

Time	User	Action
10:00 am	Paul	Config Change
9:50 am	Admin	SP Permission Added
10:10 am	Admin	SP Permission Revoked

Fraud Detection ToolBox

June 22, 2015 - Sep 21, 2015

Card #

Transaction ID	Card #	Item #	Transaction Amount	Quantity	Currency	Shipping Address	Email	Origin (IP)	Date / Time	Cancel
9198240100000	00010		50	1	USD	Wayne Jr., 82, Westwood, Florida	WayneJr@outlook.com	2.132.0.2	09/19/2015 10:10:59	✖
9198240100000	00010		20	1	GBP	19 Regent Road London, South 198, UK	latham@bt.com	2.132.0.2	09/19/2015 10:10:59	✖
9198240100000	00010		50	1	AUD	808 Broadway, North Hill, 30301, Australia	North.Hill@outlook.com	2.132.0.2	09/19/2015 10:10:59	✖
9198240100000	00010		50	1	USD	1714 South Oxford Ave, Los Angeles, CA, 90008, United States	North.Hill@outlook.com	2.132.0.2	09/19/2015 10:10:59	✖
9198240100000	00010		50	1	USD	Ballou Library, Columbia University, New York, NY 10024	latham@bt.com	172.20.21.128	09/20/2015 22:30:49	✖
9198240100000	00010		50	1	USD	Ballou Library, Columbia University, New York, NY 10024	latham@bt.com	172.20.21.128	09/20/2015 22:30:49	✖
9198240100000	00010		50	1	USD	Ballou Library, Columbia University, New York, NY 10024	latham@bt.com	172.20.21.128	09/20/2015 22:30:49	✖
9198240100000	00010		50	1	USD	Ballou Library, Columbia University, New York, NY 10024	latham@bt.com	172.20.21.128	09/20/2015 22:30:49	✖
9198240100000	00010		1000	1	USD	Ballou Library, Columbia University, New York, NY 10024	latham@bt.com	172.20.21.128	09/20/2015 22:30:49	✖
9198240100000	00010		1000	1	USD	Ballou Library, Columbia University, New York, NY 10024	latham@bt.com	172.20.21.128	09/20/2015 22:30:49	✖

1 to 10 of 10 items | 10 items per page



Fraud Detection ToolBox

June 22, 2015 - Sep 21, 2015

Card #

Transaction ID	Card #	Item #	Transaction Amount	Quantity	Currency	Shipping Address	Email	Origin (IP)	Date / Time	Cancel
9198240100000	00010		50	1	USD	Wayne Jr., 82, Westwood, Florida	WayneJr@outlook.com	2.132.0.2	09/19/2015 10:10:59	✖
9198240100000	00010		20	1	GBP	19 Regent Road London, South 198, UK	latham@bt.com	2.132.0.2	09/19/2015 10:10:59	✖
9198240100000	00010		50	1	AUD	808 Broadway, North Hill, 30301, Australia	North.Hill@outlook.com	2.132.0.2	09/19/2015 10:10:59	✖
9198240100000	00010		50	1	USD	1714 South Oxford Ave, Los Angeles, CA, 90008, United States	North.Hill@outlook.com	2.132.0.2	09/19/2015 10:10:59	✖
9198240100000	00010		50	1	USD	Ballou Library, Columbia University, New York, NY 10024	latham@bt.com	172.20.21.128	09/20/2015 22:30:49	✖
9198240100000	00010		50	1	USD	Ballou Library, Columbia University, New York, NY 10024	latham@bt.com	172.20.21.128	09/20/2015 22:30:49	✖
9198240100000	00010		50	1	USD	Ballou Library, Columbia University, New York, NY 10024	latham@bt.com	172.20.21.128	09/20/2015 22:30:49	✖
9198240100000	00010		50	1	USD	Ballou Library, Columbia University, New York, NY 10024	latham@bt.com	172.20.21.128	09/20/2015 22:30:49	✖
9198240100000	00010		1000	1	USD	Ballou Library, Columbia University, New York, NY 10024	latham@bt.com	172.20.21.128	09/20/2015 22:30:49	✖
9198240100000	00010		1000	1	USD	Ballou Library, Columbia University, New York, NY 10024	latham@bt.com	172.20.21.128	09/20/2015 22:30:49	✖

1 to 10 of 10 items | 10 items per page



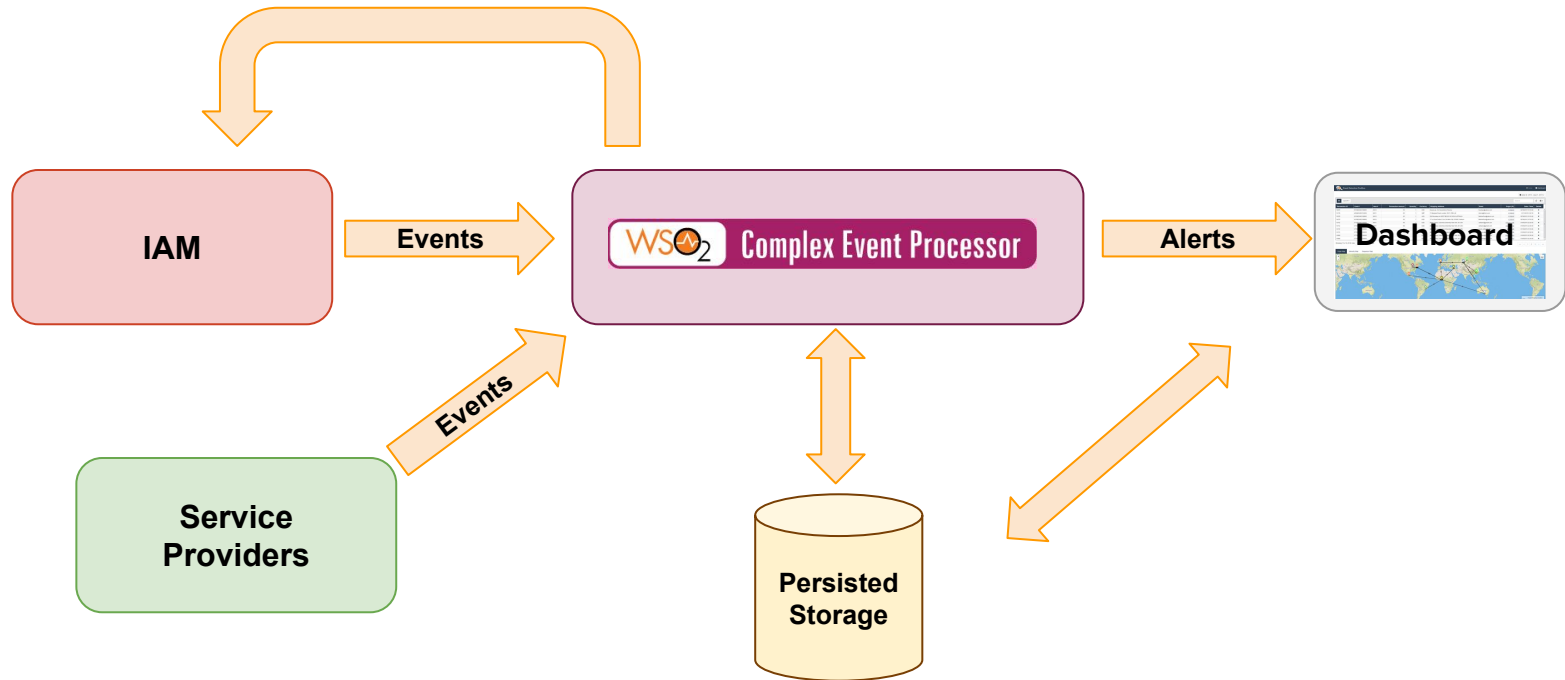
Investigate

Access historical data using

- Expressive Querying
- Easy Filtering
- Useful Visualizations

to isolate incidents and unearth relationships

Deployment

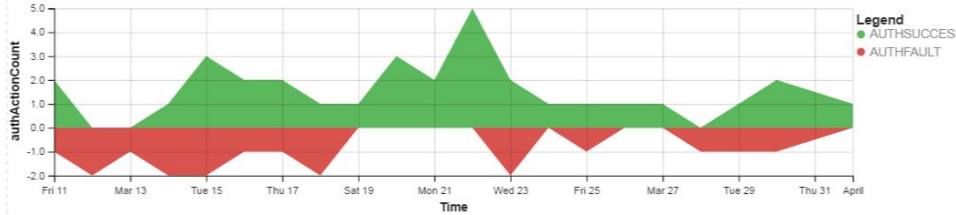


Challenges

Unusual behaviour?

Last Hour
Last 24 Hours
Last 30 Days
Last Year
Custom

LOGIN ATTEMPTS OVER TIME



Total login attempts
50

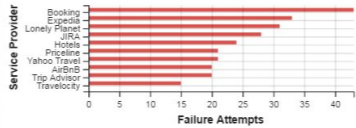
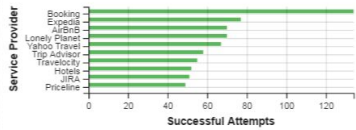


Success Rate

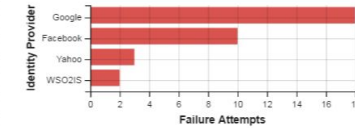
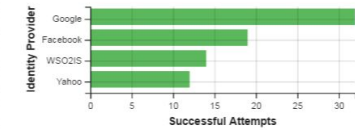


Failure Rate

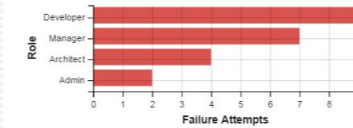
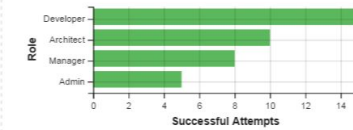
TOP SERVICE PROVIDERS



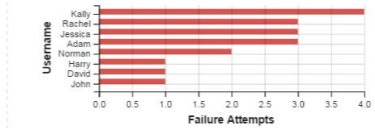
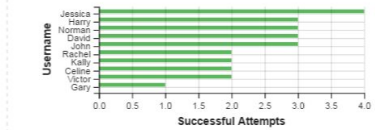
TOP IDENTITY PROVIDERS



TOP ROLES



TOP USERS



Big Data Challenge

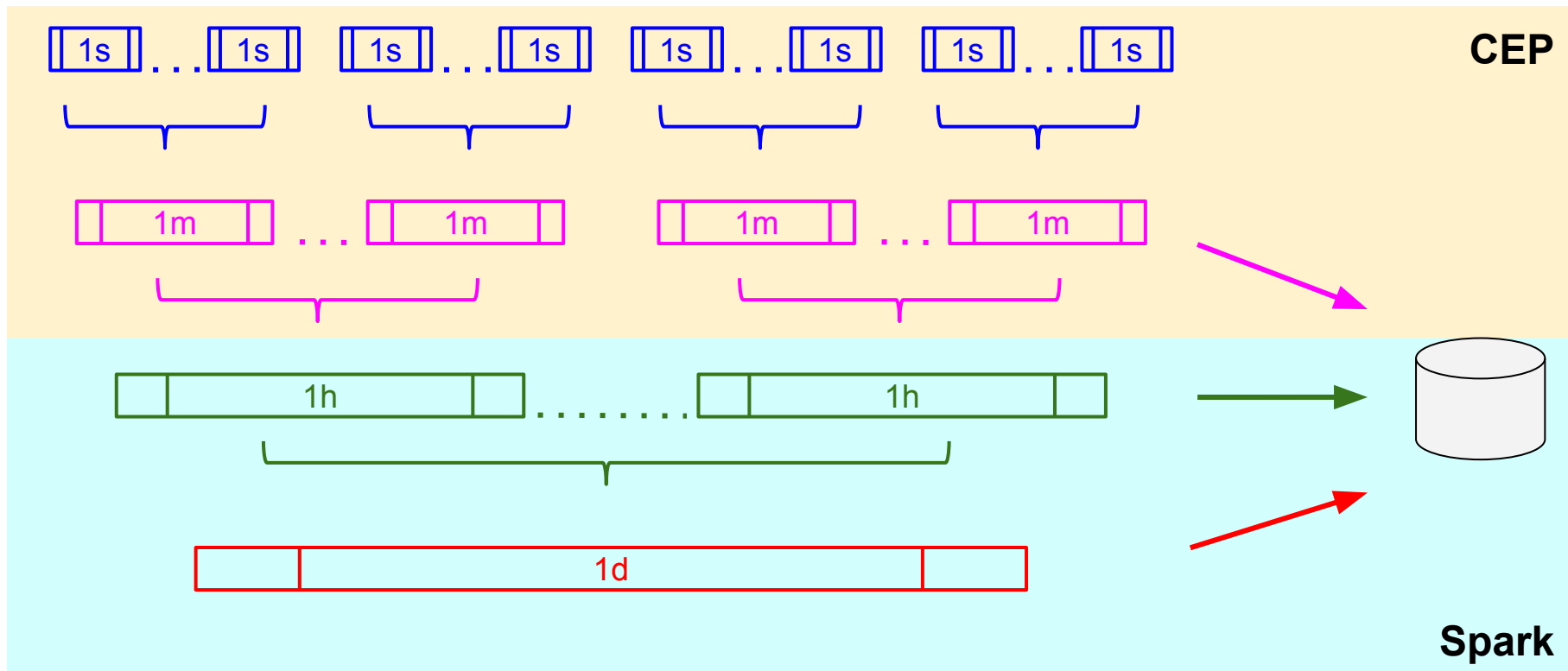
- Millions of Events
- Highly Dimensional

EventID	Timestamp	Auth Success	Username	Roles	Service Provider	IDP	IP
1	1420092114000	True	Norman	Dev; Admin	Expedia	Google	100.3.2.88
2	1420092114200	True	John	Dev	Concur	Facebook	10.13.2.15
3	1420092115500	False	Mary	QA	Ebay	Facebook	20.3.2.132

- Real-time Dashboards

Last Hour	Last 24 Hours	Last 30 Days	Last Year	 Custom ▾
-----------	---------------	--------------	-----------	--------------------------------------------------------------------------------------------

Fight against Time



Siddhi & Spark

```
from AuthEventStream#window.TimeBatch(1 sec)
select sum(AuthCount), year, month, date, hour, min, sec
insert into PerSecAuthCountStream
```

Siddhi

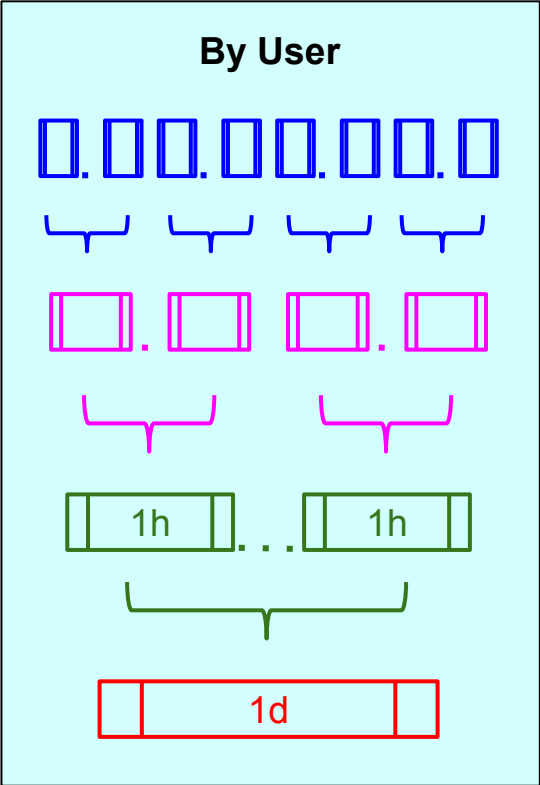
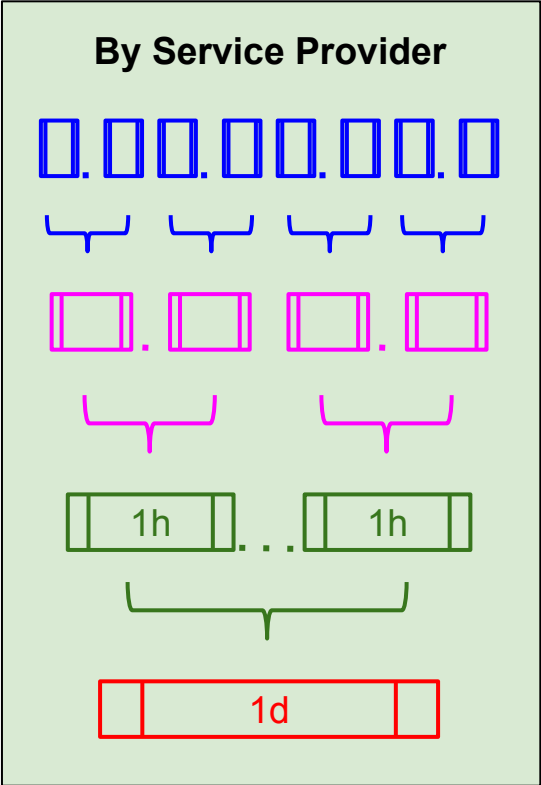
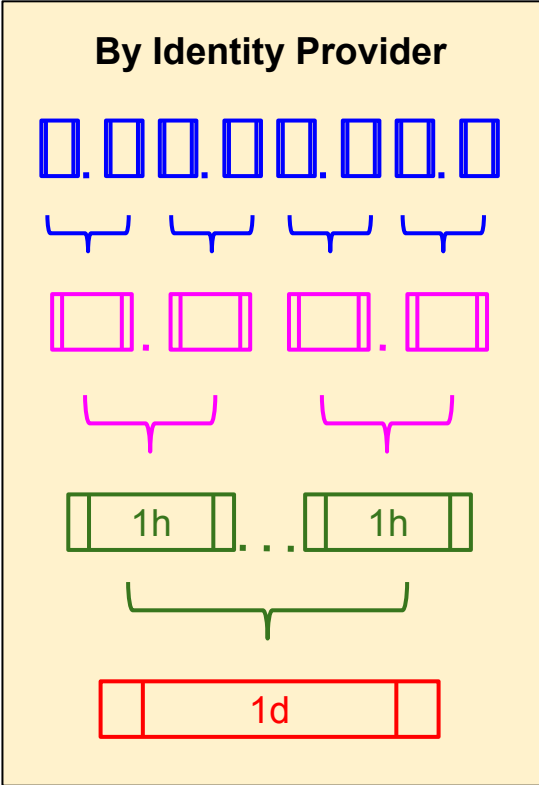
```
from PerSecAuthCountStream#window.TimeBatch(1 min)
select sum(AuthCount), year, month, date, hour, min
insert into PerMinAuthCountTable
```

```
insert into PerHourAuthCountTable
select sum(AuthCount), year, month, date, hour
from PerMinAuthCountTable
group by year, month, date, hour
```

```
insert into PerDayAuthCountTable
select sum(AuthCount), year, month, date
from PerHourAuthCountTable
group by year, month, date
```

Spark

Battling Dimensionality



North America



Europe



Middle East and Asia-Pacific



South America



Contact us !

