

---

# Azure Stream Analytics Project: On-demand real-time analytics

---

Efstratios Gounidellis  
stratos.gounidellis [at] gmail.com

Lamprini Koutsokera  
lkoutsokera [at] gmail.com

Course: "Big Data Management Systems"  
Professor: Damianos Chatziantoniou

Department of Management Science & Technology

School of Business  
Athens University of Economics & Business

May 26, 2017

## Contents

<b>1</b>	<b>Introduction</b>	<b>3</b>
<b>2</b>	<b>Azure Stream Analytics: Configuration</b>	<b>3</b>
<b>3</b>	<b>Azure Stream Analytics: Queries</b>	<b>3</b>

# 1 Introduction

This assignment is a part of a project implemented in the context of the course "Big Data Management Systems" taught by Prof. Chatziantoniou in the Department of Management Science and Technology (AUEB). The aim of the project is to familiarize the students with big data management systems such as Hadoop, Redis, MongoDB and Neo4j.

In the context of this assignment on Stream Analytics, Azure Stream Analytics will be used in order to process a data stream of ATM transactions and answer stream queries. The schema of the stream is the following: (ATMCode, CardNumber, Type, Amount).

## 2 Azure Stream Analytics: Configuration

In order to execute stream processes and queries on Azure Stream Analytics platform the following steps are required:

1. Create an Azure account.
2. Setup an Event Hub.
3. Feed the Event Hub with data.
4. Setup a Storage account.
5. Upload the Reference Data files (if any).
6. Create a Blob Storage Output.
7. Setup a Stream Analytics Job.
8. Use the Event Hub and/or Reference Data Files as Input.
9. Run the queries.

## 3 Azure Stream Analytics: Queries

```
1 /*
2   @author Stratos Gounidellis <stratos.gounidellis@gmail.com>
3   @author Lamprini Koutsokera <lkoutsokera@gmail.com>
4 */
5
6 /*
7   Q1: Show the total 'Amount' of 'Type = 0' transactions at 'ATM Code = 21'
8   of the last 10 minutes. Repeat as new events keep flowing
9   in (use a sliding window).
10 */
11 SELECT
12     SUM(CAST([BDSMastersInput].[Amount] AS BIGINT)) AS TotalAmount,
13     System.Timestamp AS Time
14 INTO
15     [BDSMastersOutput]
```

```

16 FROM
17     [BDSMastersInput]
18 WHERE CAST([BDSMastersInput].[Type] AS BIGINT) = 0 AND
19     CAST([BDSMastersInput].[ATMCode] AS BIGINT) = 21
20 GROUP BY SlidingWindow(minute, 10)
21
22 /*
23     Q2: Show the total 'Amount' of 'Type = 1' transactions at 'ATM Code = 21'
24     of the last hour. Repeat once every hour (use a tumbling window).
25 */
26 SELECT
27     SUM(CAST([BDSMastersInput].[Amount] AS BIGINT)) AS TotalAmount,
28     System.Timestamp AS Time
29 INTO
30     [BDSMastersOutput]
31 FROM
32     [BDSMastersInput]
33 WHERE CAST([BDSMastersInput].[Type] AS BIGINT) = 1 AND
34     CAST([BDSMastersInput].[ATMCode] AS BIGINT) = 21
35 GROUP BY TumblingWindow(hour, 1)
36
37 /*
38     Q3: Show the total 'Amount' of 'Type = 1' transactions at 'ATM Code = 21'
39     of the last hour. Repeat once every 30 minutes (use a hopping window).
40 */
41 SELECT
42     SUM(CAST([BDSMastersInput].[Amount] AS BIGINT)) AS TotalAmount,
43     System.Timestamp AS Time
44 INTO
45     [BDSMastersOutput]
46 FROM
47     [BDSMastersInput]
48 WHERE CAST([BDSMastersInput].[Type] AS BIGINT) = 1 AND
49     CAST([BDSMastersInput].[ATMCode] AS BIGINT) = 21
50 GROUP BY HoppingWindow(minute, 60, 30)
51
52 /*
53     Q4: Show the total 'Amount' of 'Type = 1' transactions per 'ATM Code'
54     of the last one hour (use a sliding window).
55 */
56 SELECT
57     CAST([BDSMastersInput].[ATMCode] AS BIGINT) AS AtmCode,
58     SUM(CAST([BDSMastersInput].[Amount] AS BIGINT)) AS TotalAmount,
59     System.Timestamp AS Time
60 INTO
61     [BDSMastersOutput]
62 FROM
63     [BDSMastersInput]
64 WHERE CAST([BDSMastersInput].[Type] AS BIGINT) = 1
65 GROUP BY CAST([BDSMastersInput].[ATMCode] AS BIGINT),
66     SlidingWindow(hour, 1)
67
68 /*
69     Q5: Show the total 'Amount' of 'Type = 1' transactions per 'Area Code'

```

```

70     of the last hour. Repeat once every hour (use a tumbling window).
71 */
72 SELECT
73     CAST([atmRef].[area_code] AS BIGINT) AS AreaCode,
74     SUM(CAST([BDSMastersInput].[Amount] AS BIGINT)) AS TotalAmount,
75     System.Timestamp AS Time
76 INTO
77     [BDSMastersOutput]
78 FROM
79     [BDSMastersInput]
80 INNER JOIN [atmRef]
81     ON CAST([atmRef].[atm_code] AS BIGINT) = CAST([BDSMastersInput].[atmCode]
82     AS BIGINT)
83 WHERE CAST([BDSMastersInput].[Type] AS BIGINT) = 1
84 GROUP BY CAST([atmRef].[area_code] AS BIGINT),
85     TumblingWindow(hour, 1)
86 /*
87     Q6: Show the total 'Amount' per ATM's 'City' and Customer's 'Gender'
88     of the last hour. Repeat once every hour (use a tumbling window).
89 */
90 SELECT
91     [areaRef].[area_city] AS City,
92     [customerRef].[gender] AS Gender,
93     SUM(CAST([BDSMastersInput].[Amount] AS BIGINT)) AS TotalAmount,
94     System.Timestamp AS Time
95 INTO
96     [BDSMastersOutput]
97 FROM
98     [BDSMastersInput]
99 INNER JOIN [customerRef]
100    ON CAST([customerRef].[card_number] AS BIGINT) = CAST([BDSMastersInput].[
101    CardNumber] AS BIGINT)
102 INNER JOIN [atmRef]
103    ON CAST([atmRef].[atm_code] AS BIGINT) = CAST([BDSMastersInput].[ATMCode]
104    AS BIGINT)
105 INNER JOIN [areaRef]
106    ON CAST([areaRef].[area_code] AS BIGINT) = CAST([atmRef].[area_code] AS
107    BIGINT)
108 GROUP BY [areaRef].[area_city],
109     [customerRef].[gender],
110     TumblingWindow(hour, 1)
111 /*
112     Q7: Alert (SELECT '1') if a Customer has performed two transactions
113     of 'Type = 1' in a window of an hour (use a sliding window).
114 */
115 SELECT
116     [customerRef].[first_name] AS Name,
117     [customerRef].[last_name] AS Surname,
118     CAST([BDSMastersInput].[CardNumber] AS BIGINT) AS CardNo,
119     COUNT (*) AS Transactions,
120     System.Timestamp AS Time
121 INTO

```

```

120     [BDSMastersOutput]
121 FROM
122     [BDSMastersInput]
123 INNER JOIN [customerRef]
124     ON CAST([customerRef].[card_number] AS BIGINT) = CAST([BDSMastersInput].[
    CardNumber] AS BIGINT)
125 WHERE CAST([BDSMastersInput].[Type] AS BIGINT) = 1
126 GROUP BY [customerRef].[first_name],
127     [customerRef].[last_name],
128     CAST([BDSMastersInput].[CardNumber] AS BIGINT),
129     SlidingWindow(hour, 1)
130 HAVING Transactions = 2
131
132 /*
133     Q8: Alert (SELECT '1') if the 'Area Code' of the ATM of the transaction
134     is not the same as the 'Area Code' of the 'Card Number'
135     (Customer's Area Code) - (use a sliding window).
136 */
137 SELECT
138     CAST([atmRef].[area_code] AS BIGINT) AS AtmAreaCode,
139     CAST([customerRef].[area_code] AS BIGINT) AS CustomerAreaCode,
140     COUNT (*),
141     System.Timestamp AS Time
142 INTO
143     [BDSMastersOutput]
144 FROM
145     [BDSMastersInput]
146 INNER JOIN [customerRef]
147     ON CAST([customerRef].[card_number] AS BIGINT) = CAST([BDSMastersInput].[
    CardNumber] AS BIGINT)
148 INNER JOIN [atmRef]
149     ON CAST([atmRef].[atm_code] AS BIGINT) = CAST([BDSMastersInput].[ATMCode]
    AS BIGINT)
150 WHERE CAST([atmRef].[area_code] AS BIGINT) != CAST([customerRef].[area_code]
    AS BIGINT)
151 GROUP BY CAST([atmRef].[area_code] AS BIGINT),
152     CAST([customerRef].[area_code] AS BIGINT),
153     SlidingWindow(hour, 1)

```