

Materialized Views in a Distributed Event Stream Processing Environment

- [Home](#)
- [Objectives](#)
- [Publications](#)
- [Products](#)
- [People](#)
- [Contact](#)

Research Objectives

The main focus of this research is to define and incrementally maintain materialized views over heterogeneous structured data sources in an event stream processing environment for optimizing the execution of enterprise applications. Streams are used as a communication paradigm in this framework. Each agent, known as a Distributed Event-stream Processing Agent (DEPA), maintains or subscribes to resources, such as data and event streams as well as relational and structured XML data sources. As shown in the architectural diagram, DEPAs are capable of handling data streams from different data sources, such as application event generators, the output from continuous queries over sensor data, and streams of incremental changes from databases. Each stream has an associated schema, and is assumed to be either in a relational or XML format. The vision is that each DEPA will have a specific responsibility within the distributed application, which will increase the probability of identifying common subexpressions on which to define materialized views for improving performance.

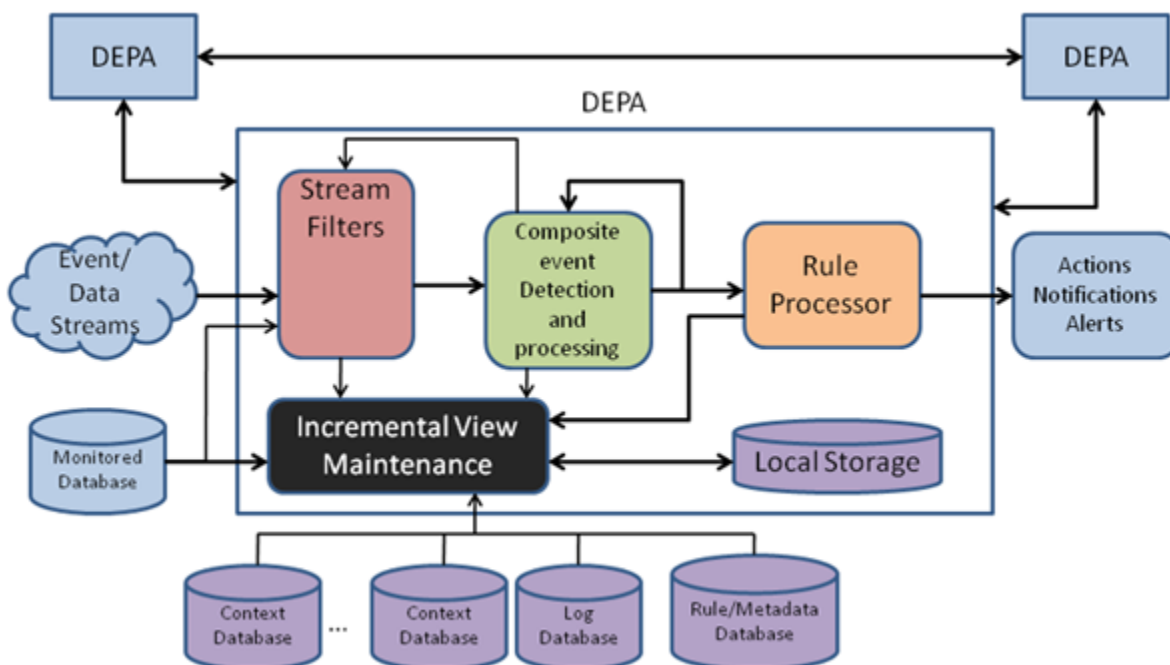


Figure 1. DEPA Architecture

The objectives of the proposed research include:

1. *Dependency analysis across different filtering queries to find common subexpressions as potential candidates for materialized partial joins.*

There are various filters used within query expressions defined over the data sources within a DEPA, such as queries over persistent data in continuous queries over streams, event specifications, and views. The dependency analysis requires the development of an integrated metadata repository that can then be used to identify common subexpressions defined within a DEPA as candidates for materialized views.

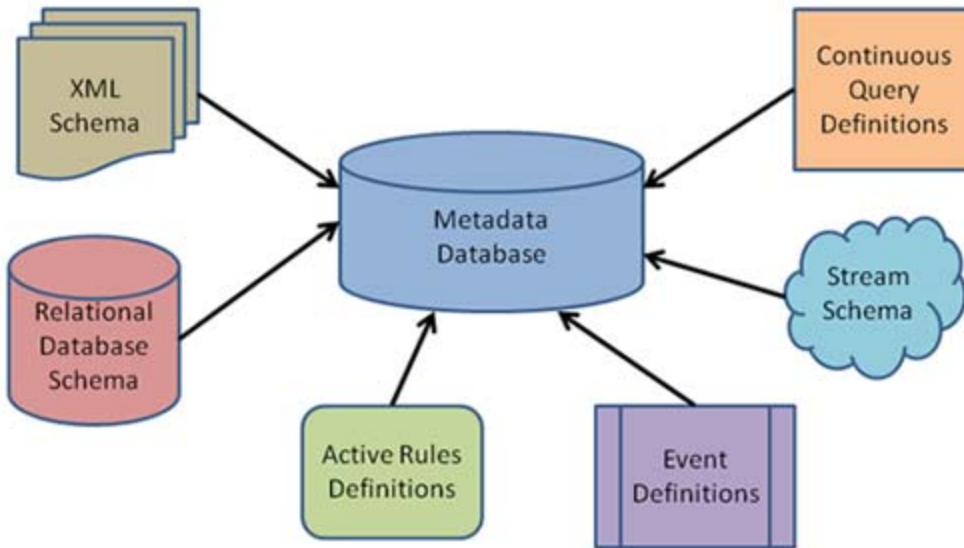


Figure 2. DEPA Metadata Repository

2. **Techniques for selectively materializing the partial joins over relational as well as XML data sources.**

The identification of which partial joins to materialize within the DEPA must be selective and include a cost-based component due to the distributed nature of the environment. Once the partial joins are identified as beneficial, the selected candidates will be defined using a materialized view definition language that supports views over relational and XML data sources.

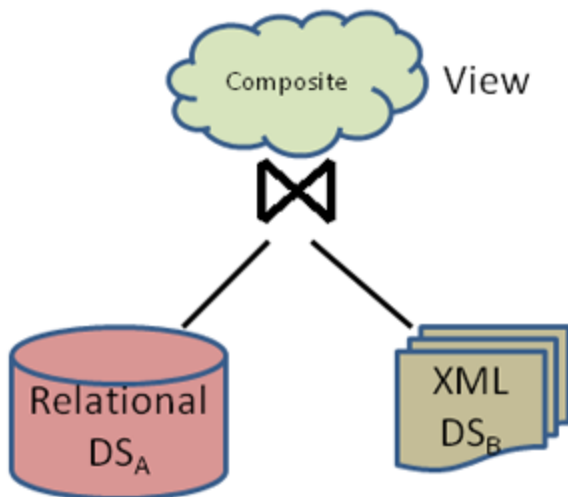


Figure 3. Composite Materialized View

3. **Incremental Evaluation of Materialized Views for Integrating Streams, Events, and Persistent Data.**

Once materialized, the views must be incrementally maintained for improved performance. Based on the event stream processing environment, the incremental evaluation techniques developed should take advantage of the deltas or changes to the data sources in their native formats as streams.

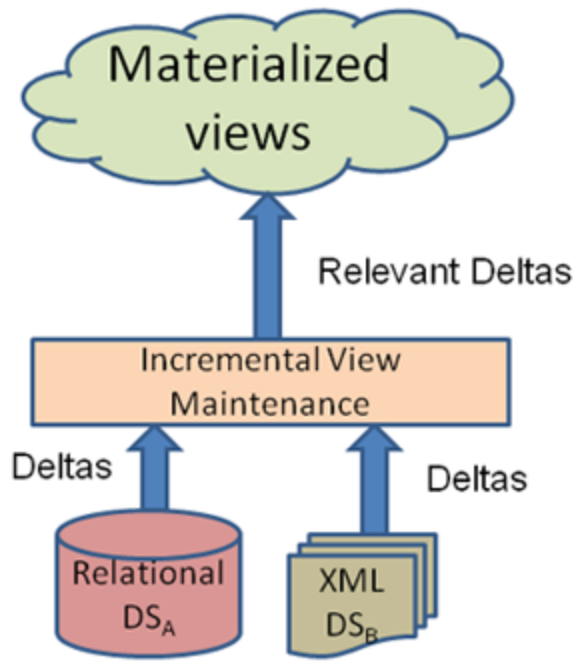


Figure 4. Incremental View Maintenance

Copyright 2009-2014. Arizona State University. All rights reserved.