BlackWhole Service Definition Document

# General Definition

BlackWhole is a data store capable of storing various schemas of data. It should be able to respond to queries for real-time data retrieval as well as async/delayed complex queries.

# Responsibility Matrix

The following table specifies which of the concerns should be addressed by the BlackWhole, and which of them are irrelevant.

|  |  |
| --- | --- |
| **YES** | * Store historical data * Enforce high-level data schema (partial schema) * Have immutable data structure (not updatable) * Allows flexible (application specific) data * Provide means for correlation tracking * Sync-result queries * Async-result queries * Allows enumeration of big data sets * Server2Server authentication * ACL * Expose unified event data stream |
| **MAYBE** | * Appendable data * Removable data (upon policy or TTL or something being fired) * Accounting * Retention policy (Define a resolution) * Digital signature validation |
| **NO** | * Full schema validation * Notification upon specific events * Graphical User Interface * Any/all *magical* inferences (i.e. full data schema, automatic segmentation, automatic tagging, etc.) |

# Instances / Variations

There might be a single instance for the BlackWhole core system and possibly multiple “adapters” each having one or more instances attached to multiple data sources (early draft).

# Communication protocol

Yet to be discussed

# Open Issues / Problems

* Who provides the “timestamp” value? (consumer or BlackWhole)
* Partial schema model
* Query language structure

# Event Model

{

TenantId: string;

AppId: string;

Timestamp: DateTime;

EventId: string;

ExternalId: string;

Tags: string[];

Data: <key,<ContentType,Content>>[] //maybe

}