These are the answers to the comments:

(2) Around 0.1 % of transfers to the Object Store fail. Since grid
    jobs run all around the world on several sites, this failures
    comprises different causes like site misconfiguration, proxy
    failures, network problems, etc.

(3) Potentially, the producer could communicate with the supervisor
    directly, but this functionality is not implemented. On the other
    hand is useful to have the CERN's ActiveMQ high availability service
    as a buffer in case the supervisor is down or overloaded.

(6) Previous system was based on HDFS CSV SeqFiles per dataset.
    EventIndex files on the Object Store had to be processed and
    transferred to HDFS by dedicated servers running Java consumers.
    Once transferred to HDFS a second step was needed to sort all
    the dataset events by EventNumber and a third step to augment
    the information with trigger names. In the new system based on
    HBASE/Phoenix there is no need to have private servers running
    the consumers, since they are replaced by Spark/Scala Loaders that
    run on the same CERN's hadoop infrastructure. Besides that, events
    are implicitly sorted by EventNumber (inside the dataset) since the
    key is built with this purpose in mind and trigger information is
    decoded by the loader itself. Furthermore, the data collection
    became more resilient because of the use of RDD (Resilient
    Distributed Dataset) by the new Spark EventIndex Loader, so
    data ingestion into HBASE has more probability to succeed in
    situations were the Object Store or HBASE have problems.