APIs or Data? An integration question.

A guide for Architects.

# Version Information

Draft, Alpha, Expect a couple of revisions soon.

Started: 8/17/23, Author: Satya Komatineni

# Background

When you want to integrate two systems, one asks should I integrate with an API or with data (by replicating, often)

## Nature of the question

The question is a bit wrong because the question is not specific.

Which integration are you talking about?

1. Intra Company
   1. System to System with specific use for that data
   2. Between two clusters of systems with multiple uses for that data
   3. Feasibility of replication due to dependencies and logic on additional data
   4. Control and variance of logic
2. Inter Company
   1. System to System with specific use for that data
   2. Between two clusters of systems with multiple uses for that data
   3. Feasibility of replication due to dependencies and logic on additional data
   4. Control and variance of logic

## Indirect Factors

1. Resilient to changing requirements.
2. Time to market
3. Ownership and Release control

## Misleading Factors

1. Premature and Excessive focus on timeliness.
2. Premature and Excessive focus on performance.
3. Short term expediency
4. Blindness due to short term view

## Fundamental Differences between API and Data outside of the obvious

1. APIs
   1. APIs tightly couple systems
   2. Inflexible in slicing/dicing of data (although some modern APIs allow this to an extent)
   3. If the consumer has a) unanticipated future needs b) multiple functions that data is needed for APIs don’t work very well
2. Data
   1. If the use of the data by the consumer is narrow this is an overkill
   2. If how the data is interpreted is a) complex or b) needs to be controlled by the producer and c) cannot be reproduced easily by the consumer, then that poses a problem for data replication

# Considerations for intra company vs inter company.

These two cases are not the same.

## Intra Company

1. Systems tend to be closer to each other.
2. Any changes to an API can be closely coordinated.
3. Availability of provider and consumer can be closely controlled.
4. Forgiving of tighter (not loose) integrations via APIs
5. However, there are cases where both APIs and Data are suitable as well in this case.
6. It is possible to construct systems in such a way that APIs can be more advantageous than data replication, but this means a cluster of systems must be “closed” to their data and only talk across the islands via APIs and that the APIs are few and not many

## Inter Company

1. Harder to anticipate future needs.
2. Data is often used for multiple systems and multiple needs.
3. An API like tighter integration makes it hard for the companies to move with flexibility.
4. Data tends to be a better option unless the integration is a) very specific and b) and data replication is not feasible or unnecessary due to smaller need.

# Specific Recommendations

## Intra Company API, When?

1. Build your systems as relatively stand-alone islands where a cluster of systems in an island operate on the same data store whereby the data doesn’t have to be replicated
2. Use APIs between the islands.
3. APIs tend to be smaller in number when implemented right.
4. Allow for paging in the API.
5. Allow for column projections in the API (More like modern entity frameworks)

## Intra Company Data, When?

1. When the consumer manipulates the data in a variety of ways
2. When the consumer has future needs that the producer don’t have to be concerned with too much
3. Data is relatively unchanging.
4. Although it is tempting to use APIs as a better solution all the time, this opportunity can often be overlooked.
5. Allows many times better multi-tenancy and isolation.

## Inter Company API, When?

1. You really want to keep the companies as minimally interlinked as possible.
2. So, consider APIs with suspicion for they will tightly couple companies.
3. When the functions are specific and limited.
4. APIs are relatively unchanging while the data changes often
5. When it is infeasible to load data due to volume or other reasons
6. When it is infeasible due to inter dependencies of data at the producer

## Inter Company Data, When?

1. When interactions with this data is extensive
2. Many systems use this data at the consumer company.
3. Keep the ownership and release cycles better.
4. Time to market considerations.
5. It is feasible to copy the data.
6. Any rules that govern the data can be replicated when looked upon as a compromise for the desired flexibility, time to market, and cost