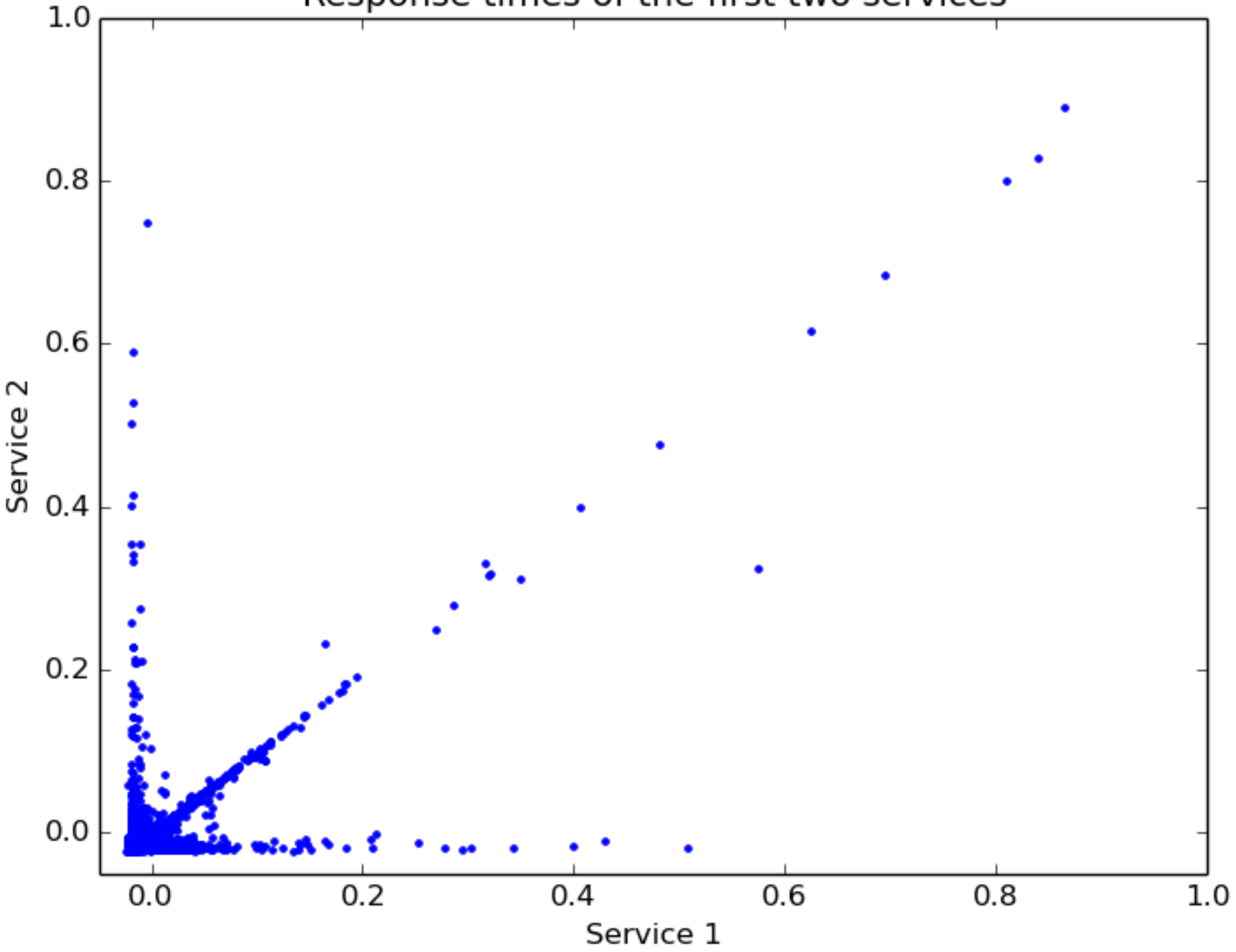
# Composite Service Status Identification on SOA based Infrastructures

# Composite Service

- Communication between applications is achieved with services.
- Composite services consist of sub-services and allow more complex operations.
- Status of a composite service is not easily identified.

# Clustering Problem

- A Composite Service relies on the Status of all of its sub-services.
- The Status of a single basic sub-service is determined by its response time.
- The response-times of the sub-services at a given time can be interpreted as a point in the phase-space.
- Multiple measurements show patterns in the phase-space.


Response times of the first two services
Service 2
Service 1
0.0
0.2
0.4
0.6
0.8
1.0

# Clustering Problem

- It is possible to attribute each point in the phase-space to a cluster corresponding to a status of the composite service.
- Necessary steps:
  - Determining the optimal number of clusters
  - Determining the actual clusters

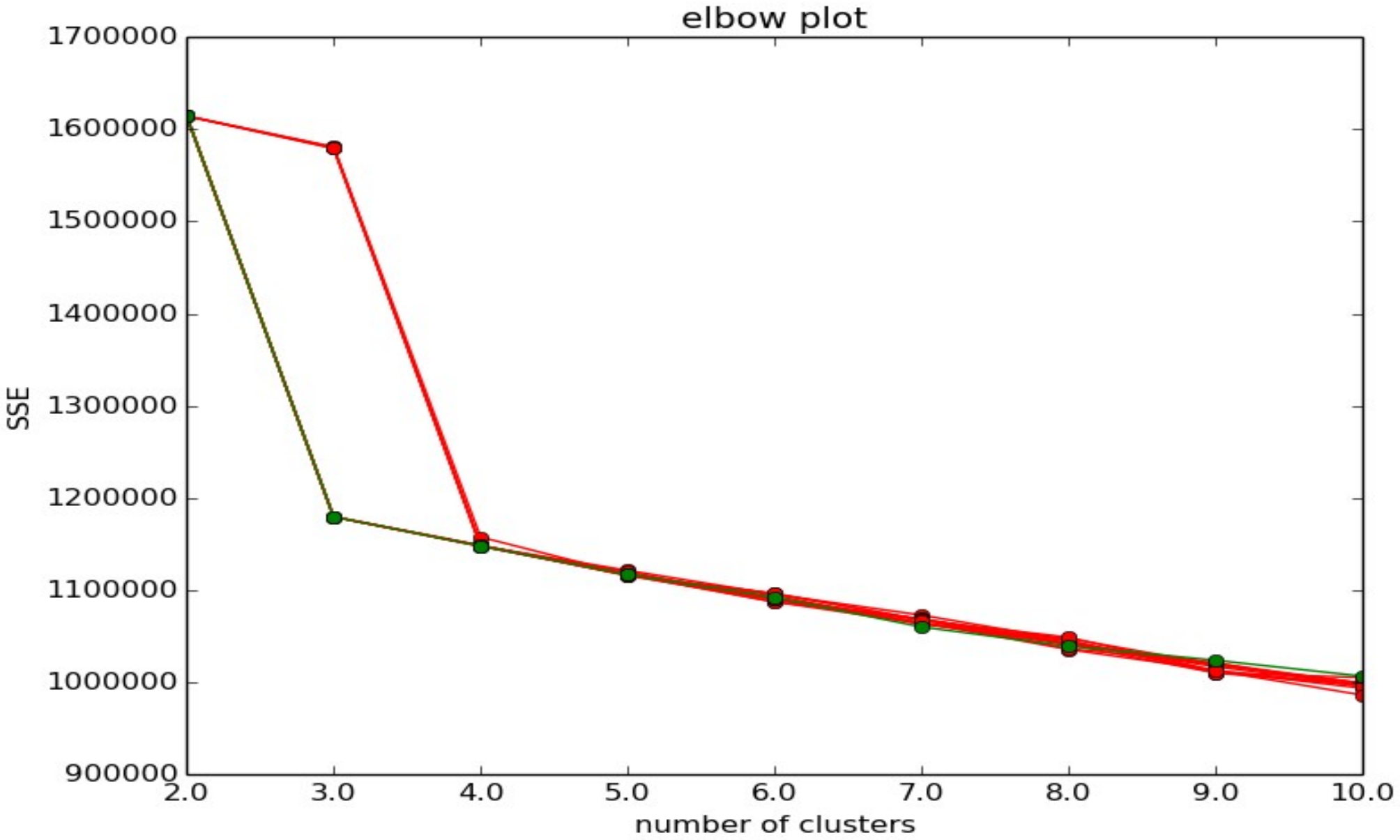
# Methods

- K-Means-clustering
  - Starts with randomly distributed cluster centres.
  - Attributes data-points to the cluster centre they are closest to.
  - Calculates new cluster centres as the mean of the points in the cluster.
- Hierarchical Clustering (Agglomerative)
  - At the beginning each data point is a separate cluster.
  - Clusters are merged based on distance.
  - Provides different levels of clustering.

# Determining the optimal number of clusters

- Using the K-Means-Algorithm for different numbers of clusters
  - The sum of squared distance (SSE) from cluster centre drops significantly at right cluster number (Elbow Method).
- Hierarchical Clustering
- Principle Component Analysis
  - Number of principle components that cover most of the variance of the data (>95%) is a good estimate for the number of clusters.

# Elbow Method Plots



# Elbow Method Plots