

# 13th International Database Engineering & Applications Symposium (IDEAS)

*Cetraro (CS) – Italy, September 16-18, 2009*

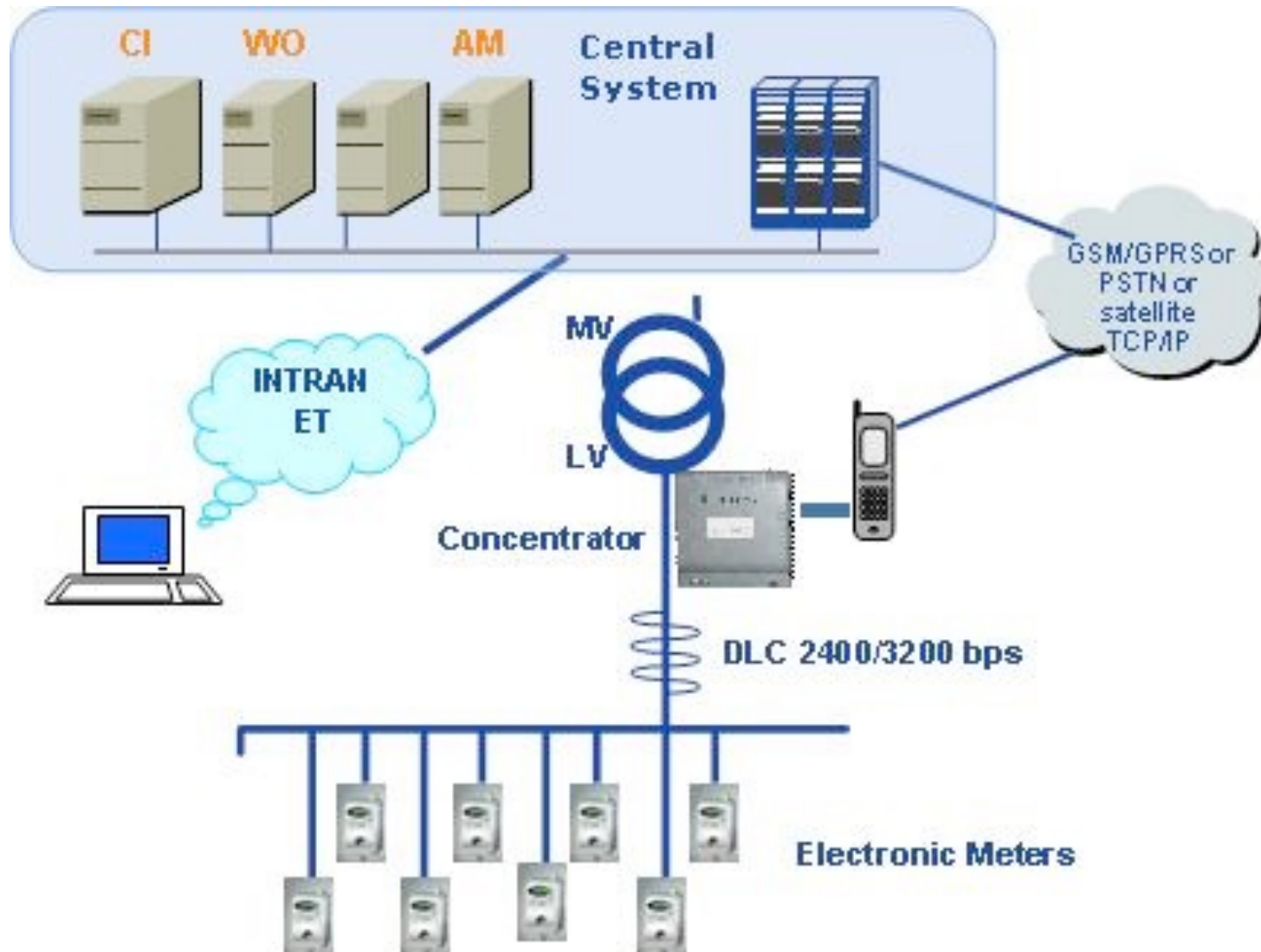
## Low-voltage Electricity Customer Profiling based on Load Data Clustering

F. Gullo, G. Ponti, A. Tagarelli (D.E.I.S. – University of Calabria)

S. Iiritano, M. Ruffolo (Exeura s.r.l.)

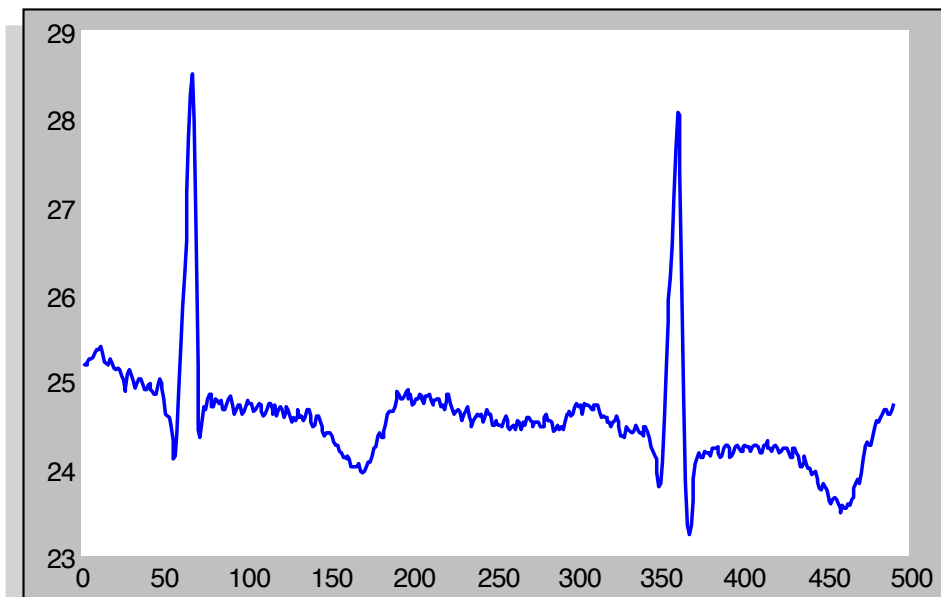
D. Labate (Enel Distribuzione s.p.a.)

# The Enel Telegestore Architecture



# Low-voltage electricity customer data

Thanks to the telegestore project, Enel is able to measure and store **load profiles** of their mass-market low-voltage (LV) customers in a flexible and effective way



**Load Profile (LP):** shape of a (group of) customer consumption chronologically ordered

# Motivations

## Main Goal:

to study the characterization of LV customers based on their consumption data

## Major emphasis:

to characterize the most typical classes of electricity customers, i.e., detecting groups of customers having similar consumption behavior

# Our proposal

- A clustering framework for electricity customer load profiles, which is supported by information on metadata (e.g., customer type, meter type, day, contract, location)
  - load profiles represented by exploiting a *time series-based* model
  - clustering algorithms: *K-Means*, *TS-Part*
  - distance measures: Euclidean, Dynamic Time Warping (DTW)

# Results

We were granted access to about 30,000 Enel Italian LV anonymous customer load profiles, measured during the period between the first week of February 2009 and the last week of March 2009

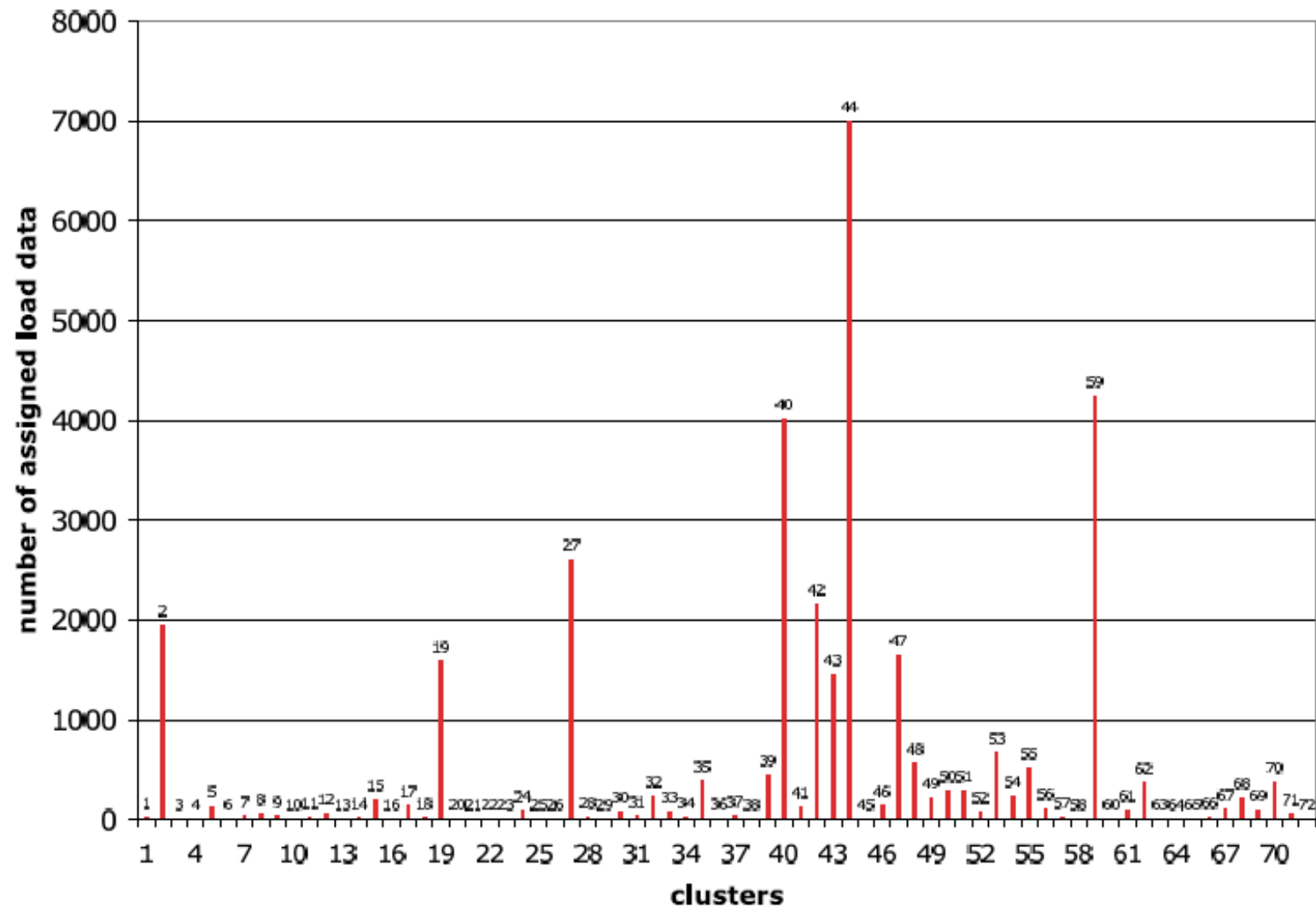
<i>clustering algorithm</i>	<i>distance measure</i>	<i># of clusters</i>	<i>MIA</i>	<i>CDI</i>
<i>K-Means</i>	Euclidean	10	9.775	0.682
<i>TS-Part</i>	Euclidean	10	9.103	0.914
<i>K-Means</i>	DTW	66	7.986	0.008
<i>TS-Part</i>	DTW	72	<b>5.514</b>	<b>0.004</b>

**Best (average) performance of clustering: Weekdays load profiles**

<i>clustering algorithm</i>	<i>distance measure</i>	<i># of clusters</i>	<i>MIA</i>	<i>CDI</i>
<i>K-means</i>	Euclidean	17	9.964	0.109
<i>TS-part</i>	Euclidean	17	6.946	0.156
<i>K-means</i>	DTW	29	13.773	0.014
<i>TS-part</i>	DTW	32	<b>11.646</b>	<b>0.012</b>

**Best (average) performance of clustering: Sundays/holidays load profiles**

# Results



**Distribution of weekdays load profiles over clusters obtained by TS-Part with DTW**