Managing Big Energy Data for (Really) Smart Grids

Torben Bach Pedersen
Daisy@CS@Aalborg University

Agenda

- **Why** Big Energy Data?
- **What** is Big Energy Data?
- **What** do we do with Big Energy Data?
  - And **how** do we do it?

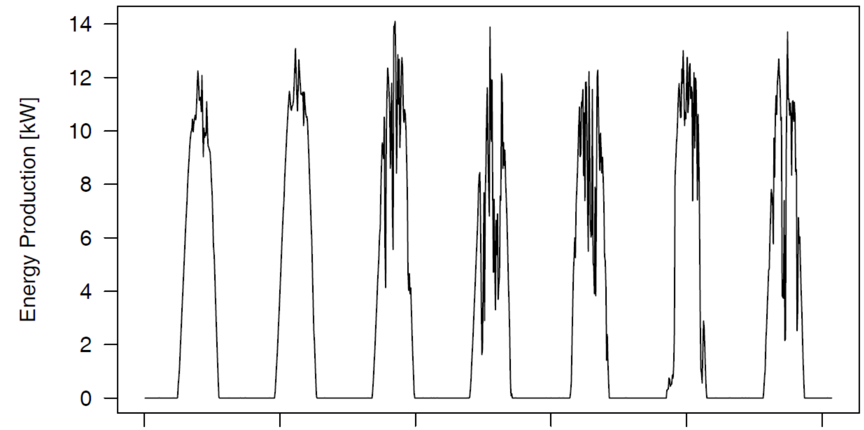
Why Big Energy Data?

- Societal challenges and solutions
  - Global warming - greenhouse gas emission cuts
  - Energy supply security – reduce energy purchased from outside
  - Nuclear risks – nuclear phaseout
- Solution: more energy from renewable energy sources
  - EU 20-20-20 goals,…
  - DK: 2020: 50% of electricity from RES, 2035: 100% electricity and heat from RES, 2050: 100% RES in **all** sectors
- Implication: move from fossile to electric energy
  - EVs and heatpumps
  - Danish electricity (not energy) consumption **tripled** in 2050

Uncertainties of Renewables

- Fluctuating Energy
  - Wind power
  - Photo Voltaic
  - Waves / Tides

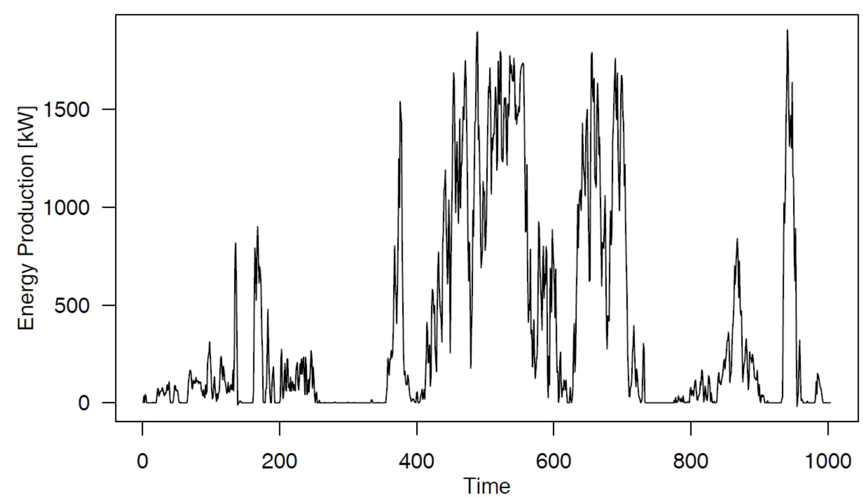
CRES
22 kWp
(Greece)



- Weather conditions determine amount of energy

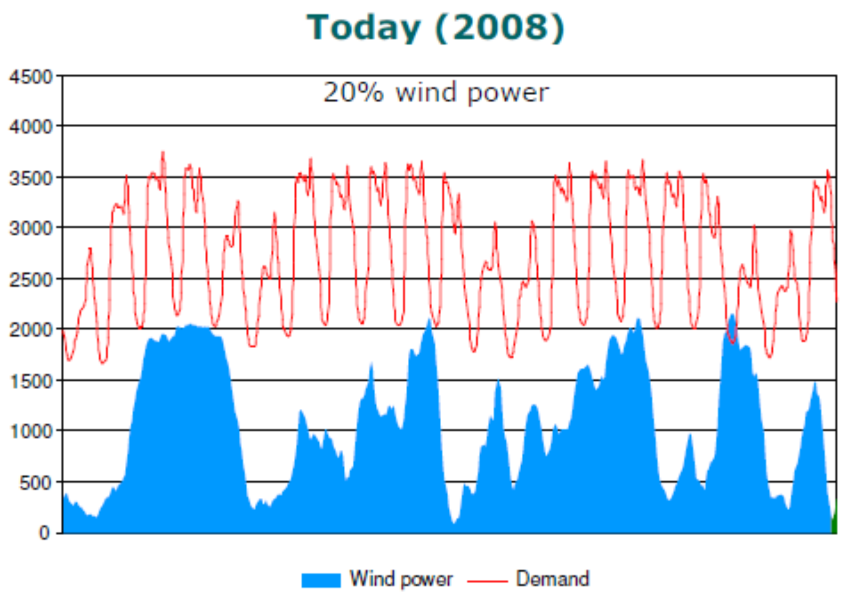
Windpark
2410 kWp
(Greece)

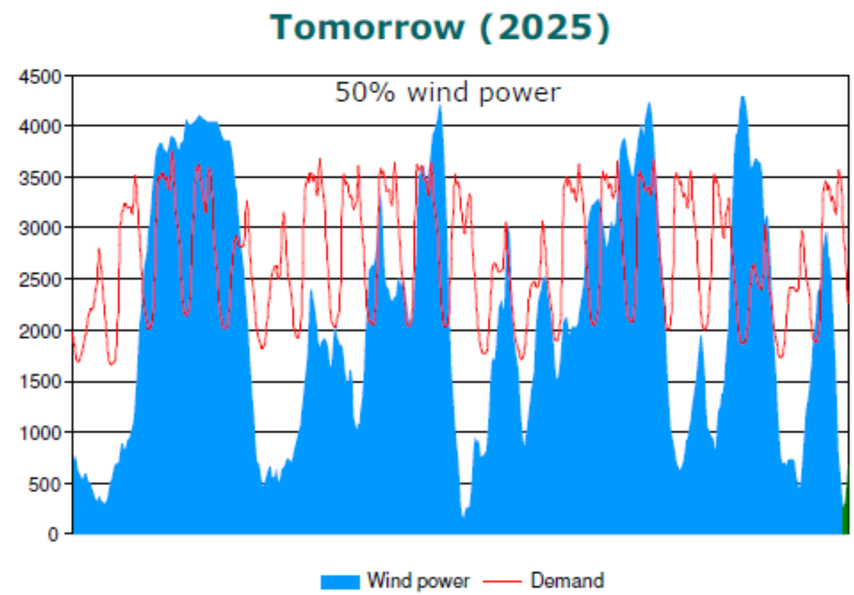
- Hydro is easy...



Too Much or Too Little Energy

2008 DK West figures





Recent DK figures for electricity produced by wind, % of total

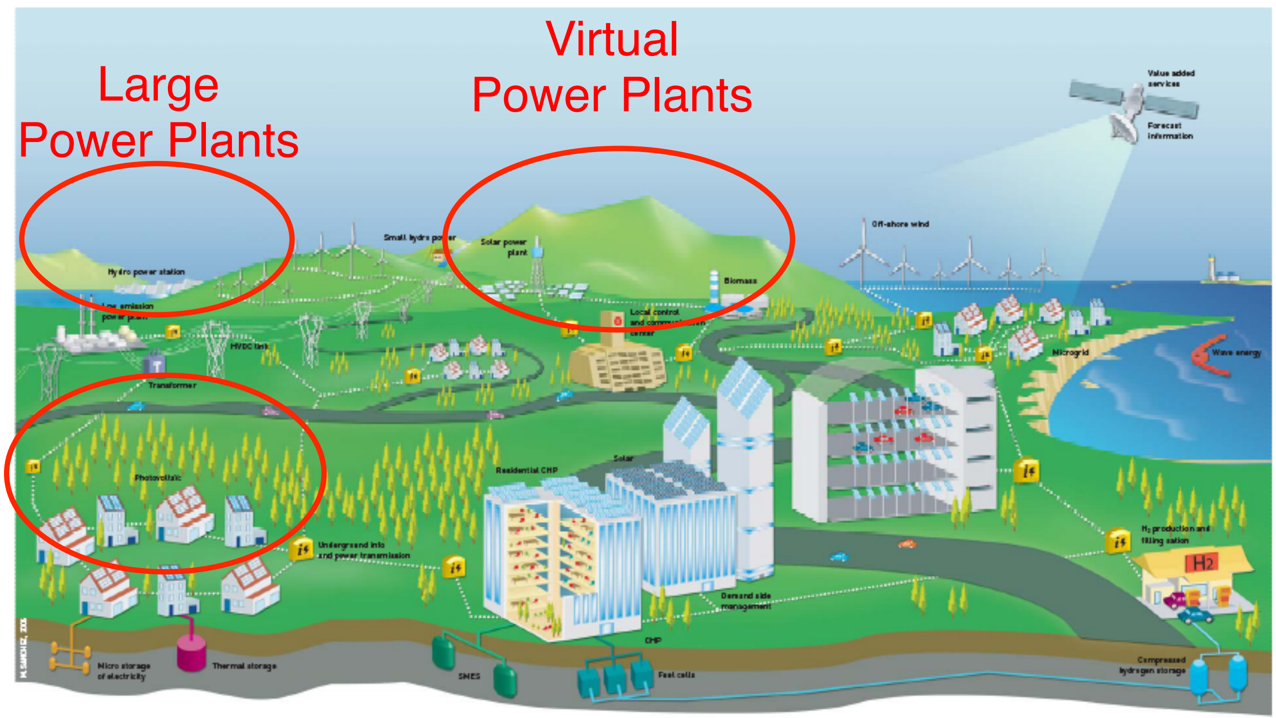
- December 2013 – 57.4%
- January-June 2014: 41.2%
- 2014: 39.1%
- The future is here today!

Flexible Demand To The Rescue

- Dishwashers and washing machines can run flexibly
- EVs can be (de-) charged flexibly during parking intervals
- Heatpumps can run flexibly within a comfort temperature interval
- Up to 80-85% of the (tripled) future demand is flexible

Future Vision: Smart Grids

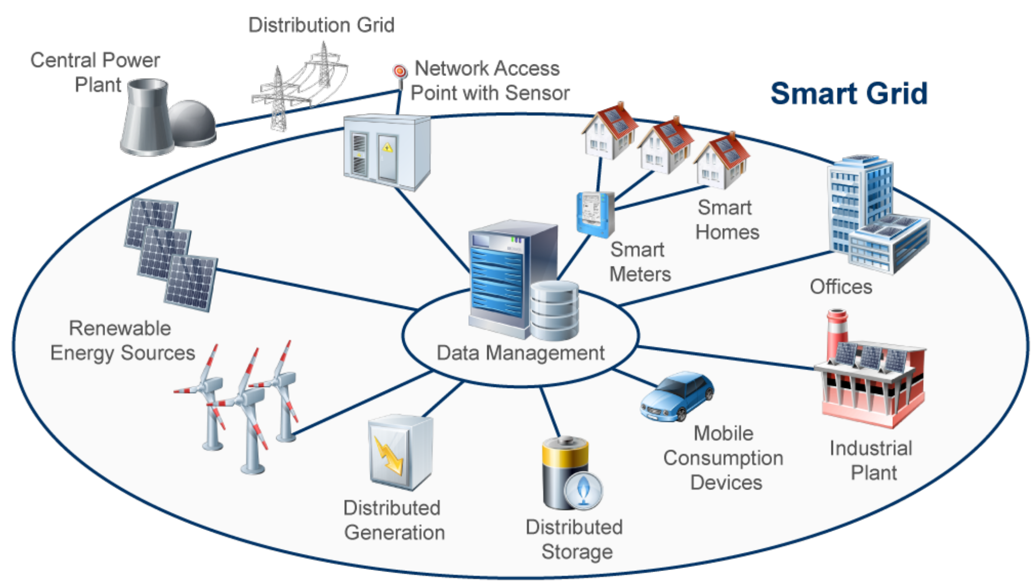
- **Smart Grids**
  - Increased flexibility of energy networks via ICT (monitor, control)
  - Goals: more RES, active customer involvement, balancing demand/supply



Smart Meter: foundation for smart grids (bi-directional communication)

Data Management Challenges

- Large-Scale Distributed Systems
  - Number of stakeholders, number of of nodes, amount of data
- High Availability / Fault Tolerance
  - Basically available, soft state, eventual consistent
- Near-Realtime Data Synchronization and Integration
  - High update rates, low latency, protocol/schema/format heterogeneity
- Advanced Analytics
  - Time series forecasting
  - Balancing
  - Classification, C
  - Clustering,
  - Association rule mining



What is Big Energy Data?

- Variety: complex data of many different types
  - Consumption data from smart meters, sockets, and appliances
  - Production data from wind, solar, power plants,…
  - Flexibility data – what demand and supply is (how) flexible?
  - Prices, weather, …
- Volume: a lot of it
  - EU consumption per prosumer per sec: 20+ trillion values/day
  - Then go to sockets/appliances and add the other data types
- Velocity: fast data
  - Real-time smart meter readings
  - So fast it hasn't even happened yet: everything is *(re-)forecasted*

- We will focus on variety today (velocity+volume tomorrow)

The MIRABEL Project

EU FP7 project (call 4)

Objective: Novel ICT Solutions for Smart Electricity Distribution Networks

mirabel-project.eu

Timeline: 01/2010 to 04/2013

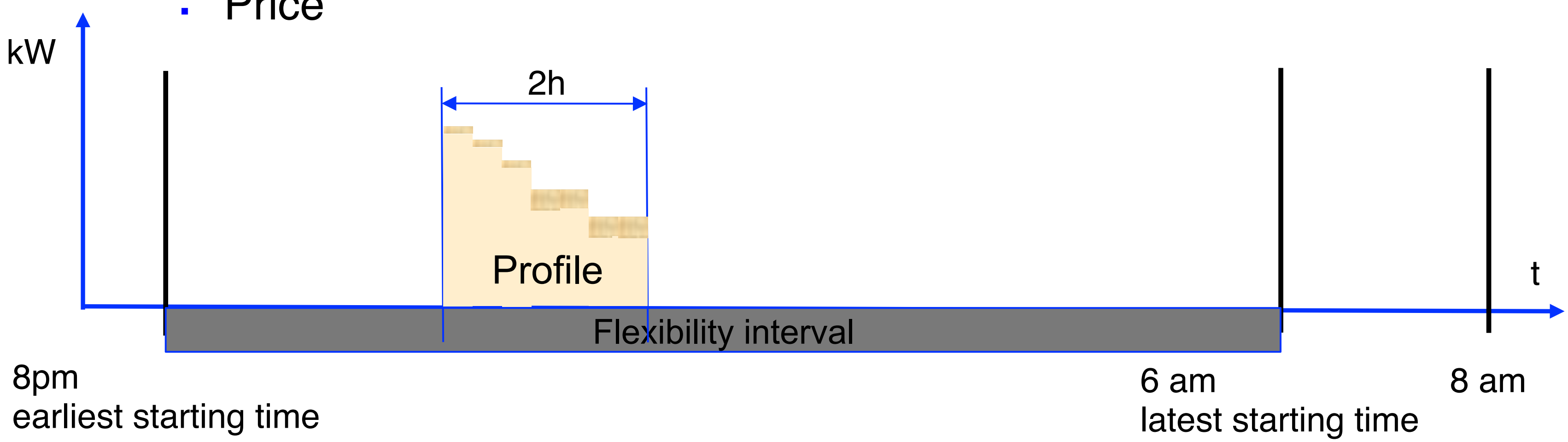
Variety: Flexibility Data

Consumers (households, SMEs,..) have some flexible, schedulable demand

- such as dishwashers, washing machines, EVs, heat pumps, …

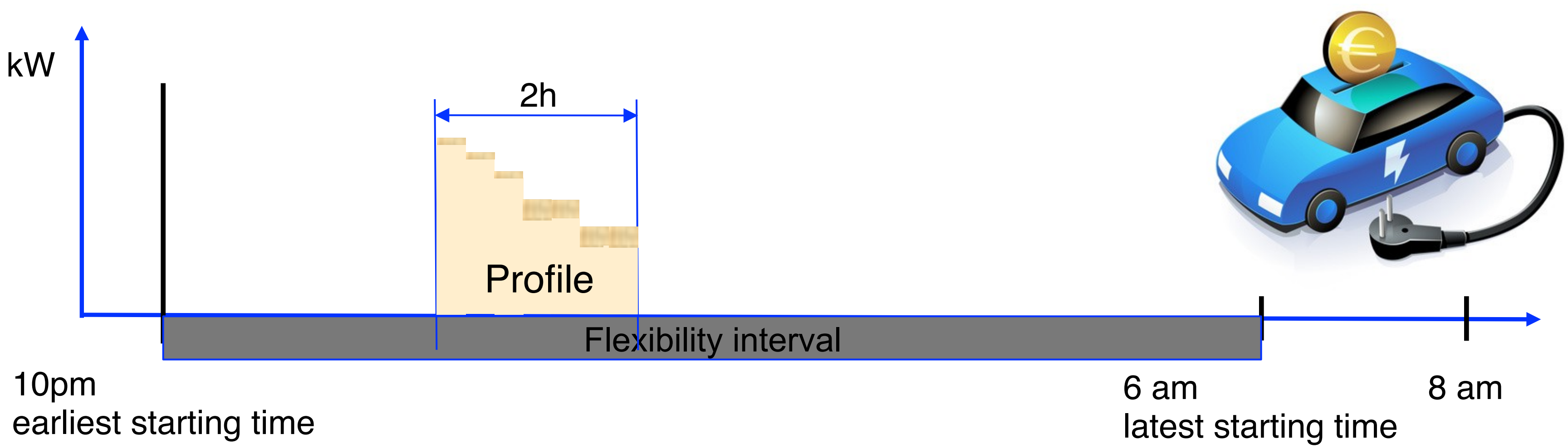
⇒ specified and treated as **flex-offers (FOs)** with **explicit** flexibility in

  - Time (flexibility interval)
  - Amount of electricity
  - Price

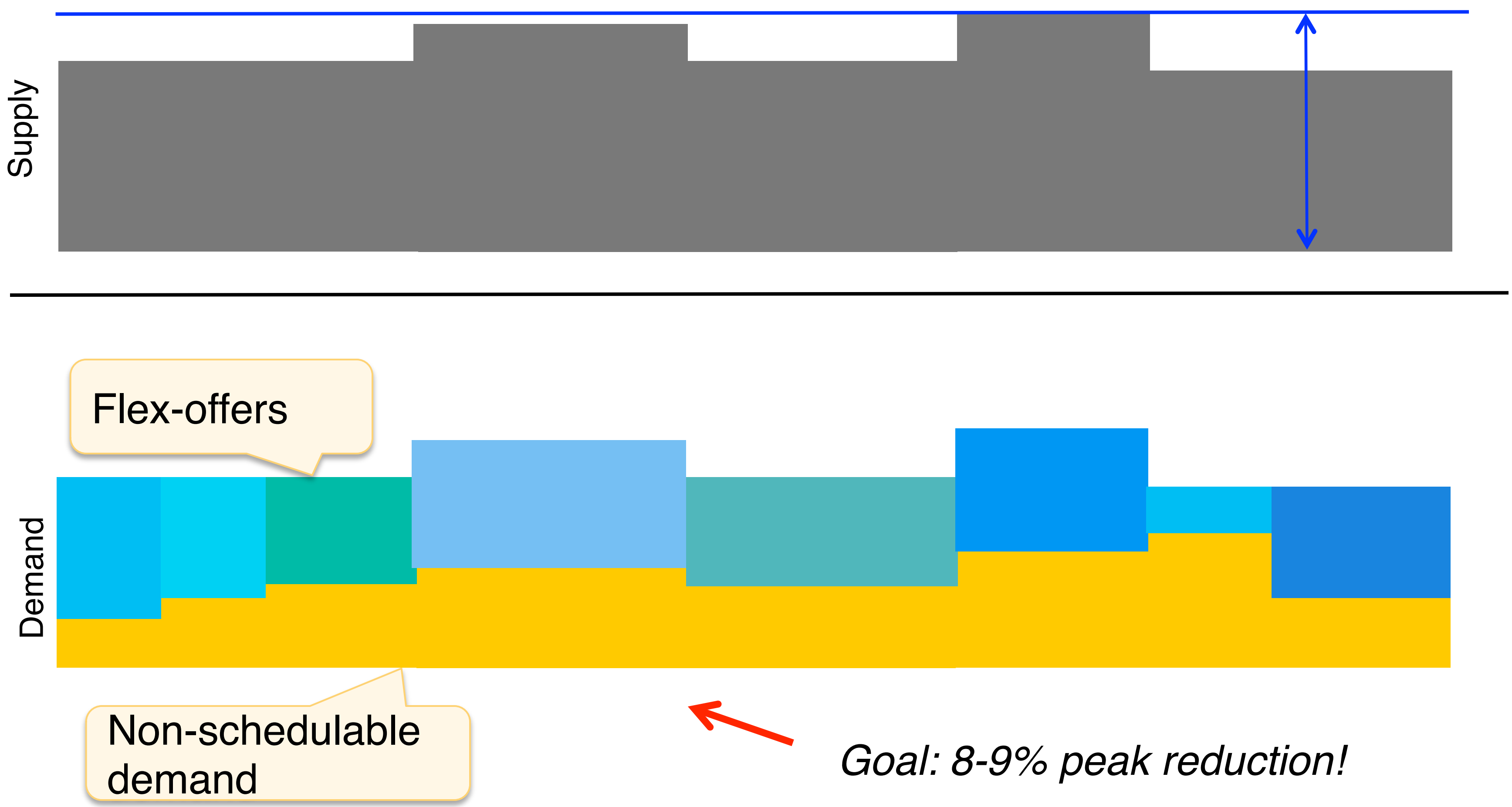


Use Case: Charging an EV

1. A consumer arrives home at 10pm and wants to recharge the electric car's battery at the lowest possible price by the next morning. Completion time is set to 6 am.
2. The prosumer node generates an FO
3. Based on weather forecasts, the trader's node schedules the FO to start energy consumption at 3am and sends back a message to the prosumer's node.
4. The consumer's node of EDMS starts supplying energy to the electric vehicle at 3am.

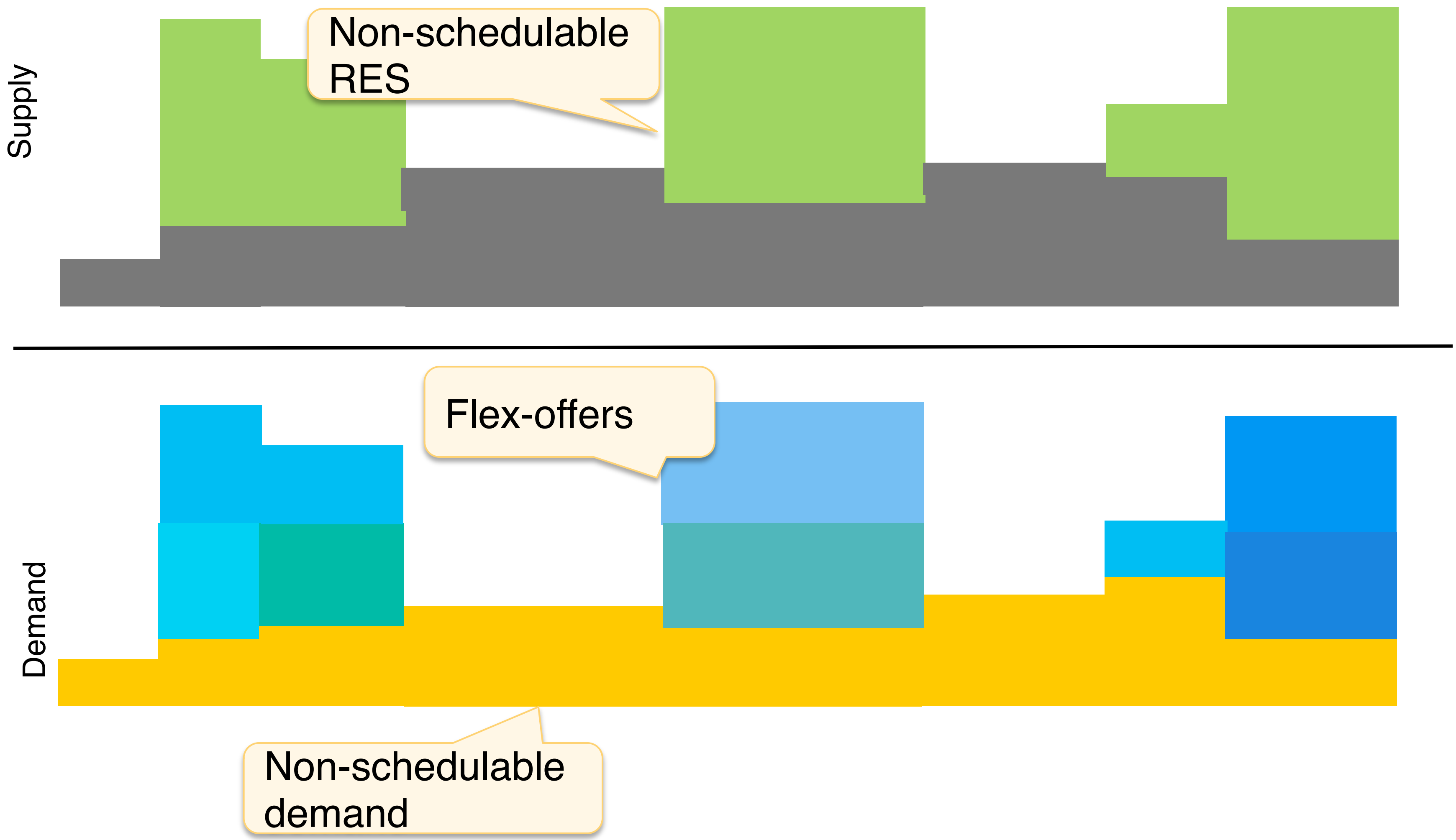


Use case: Balancing

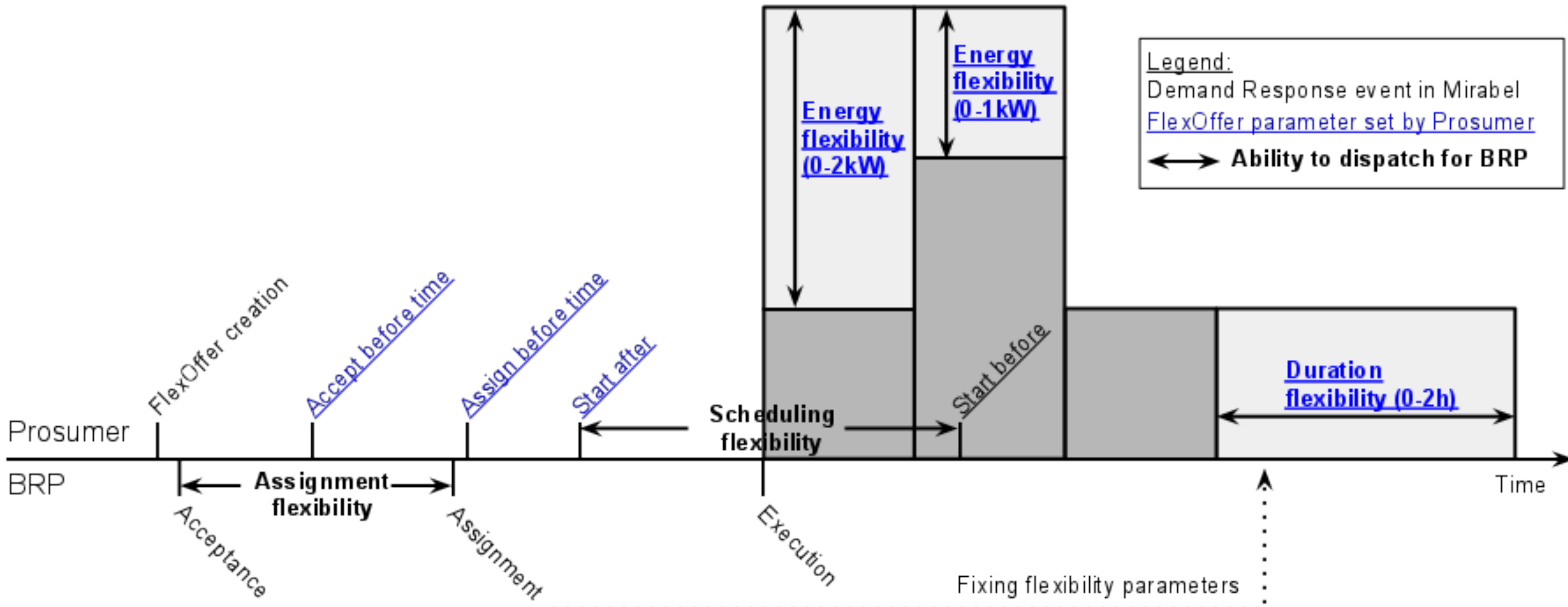


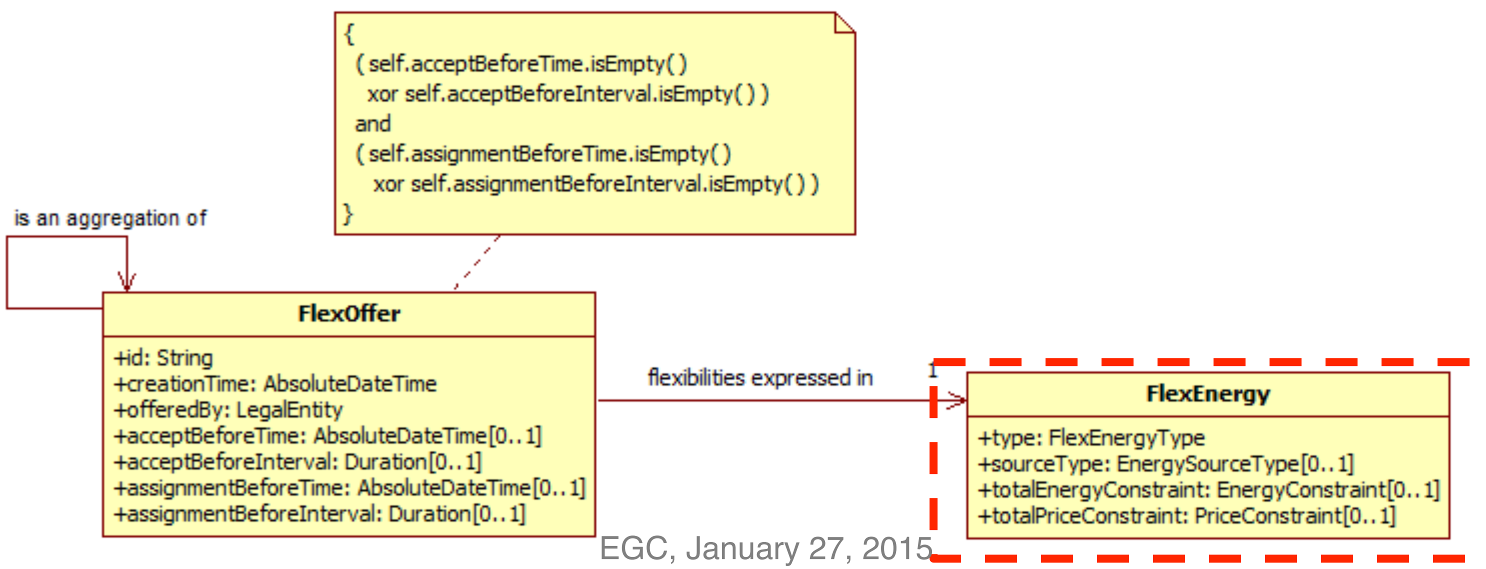
Goal: 8-9% peak reduction!

Use case: Higher rate of Renewables

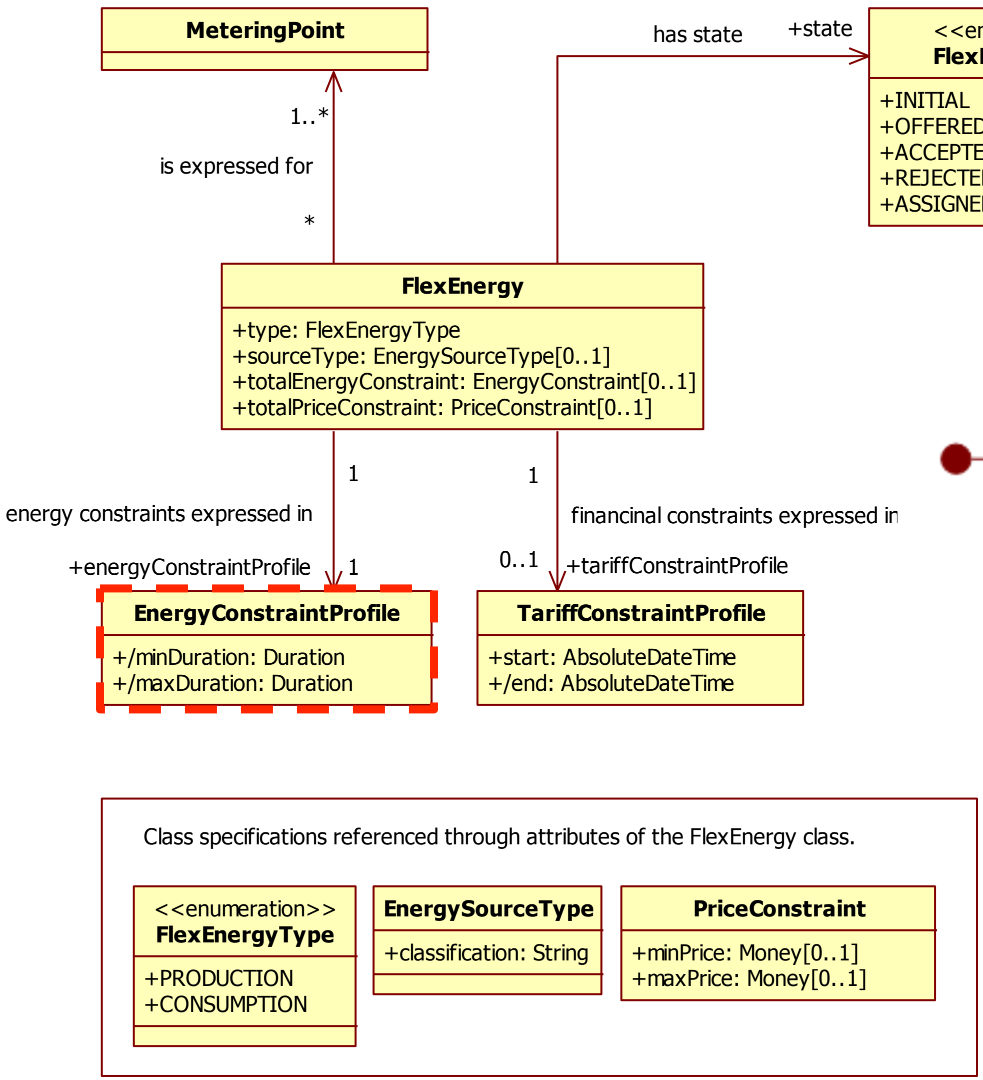


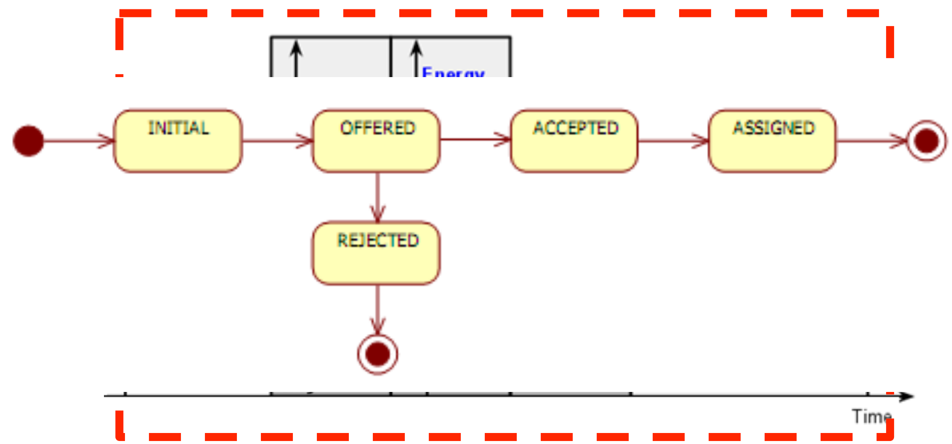
Flex-Offers: Modeling Flexibility



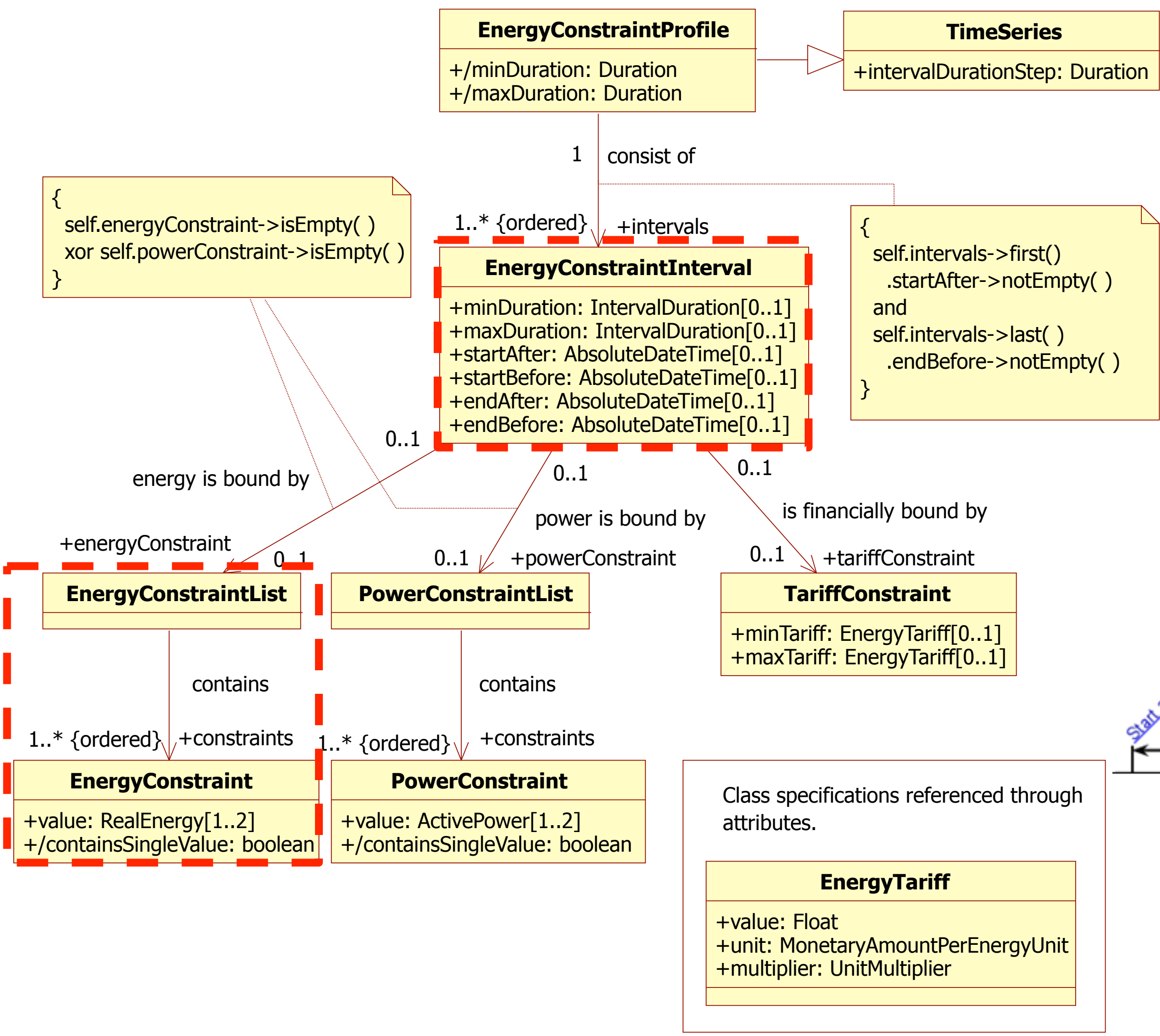


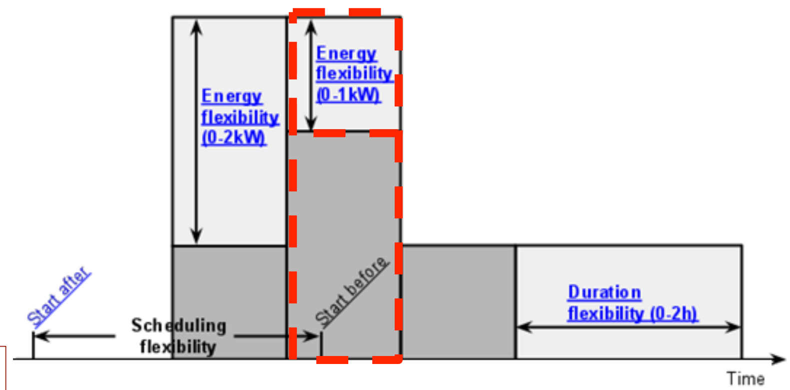
Flex-Offers: Modeling Flexibility





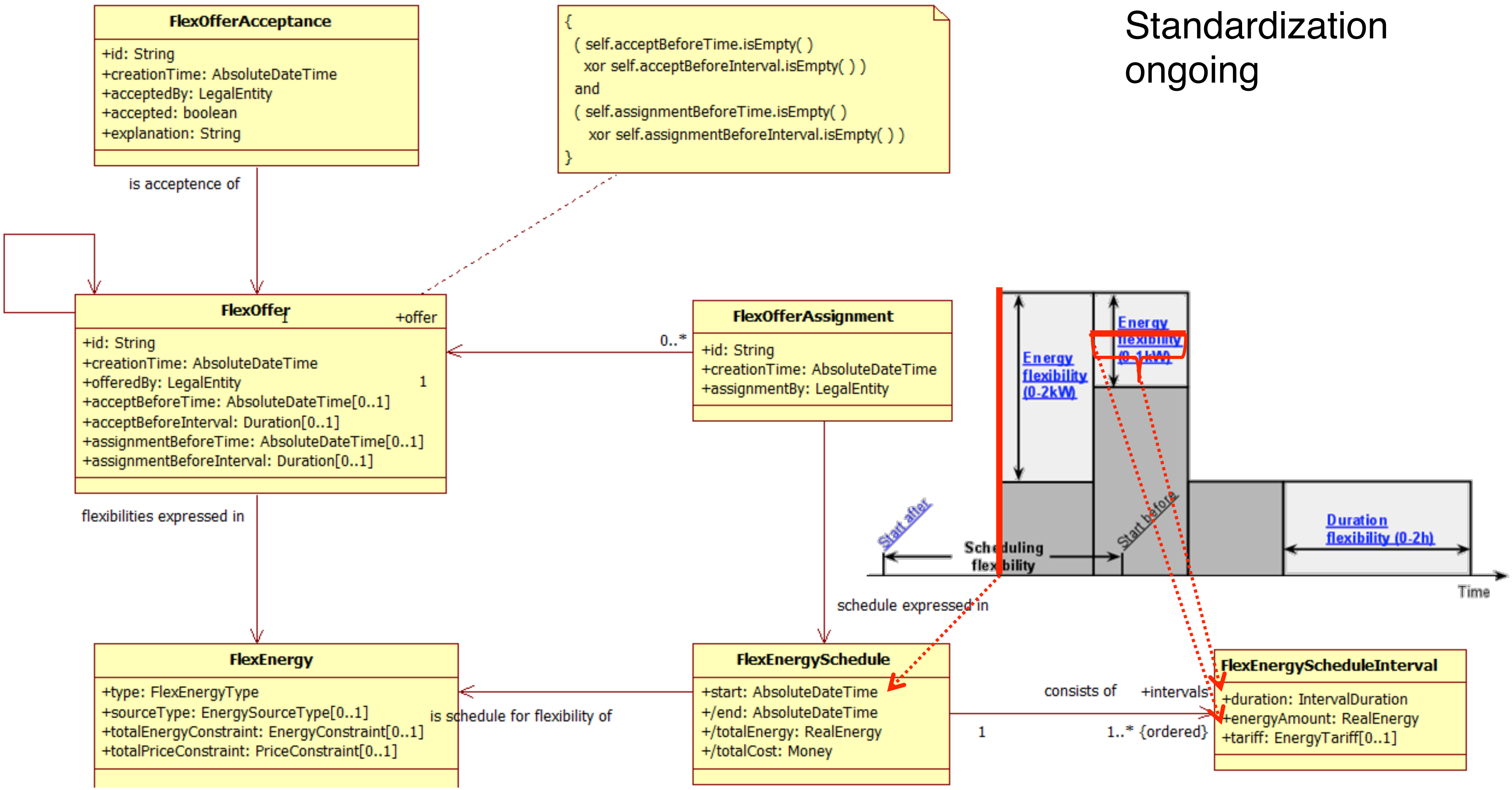
Flex-Offers: Modeling Flexibility



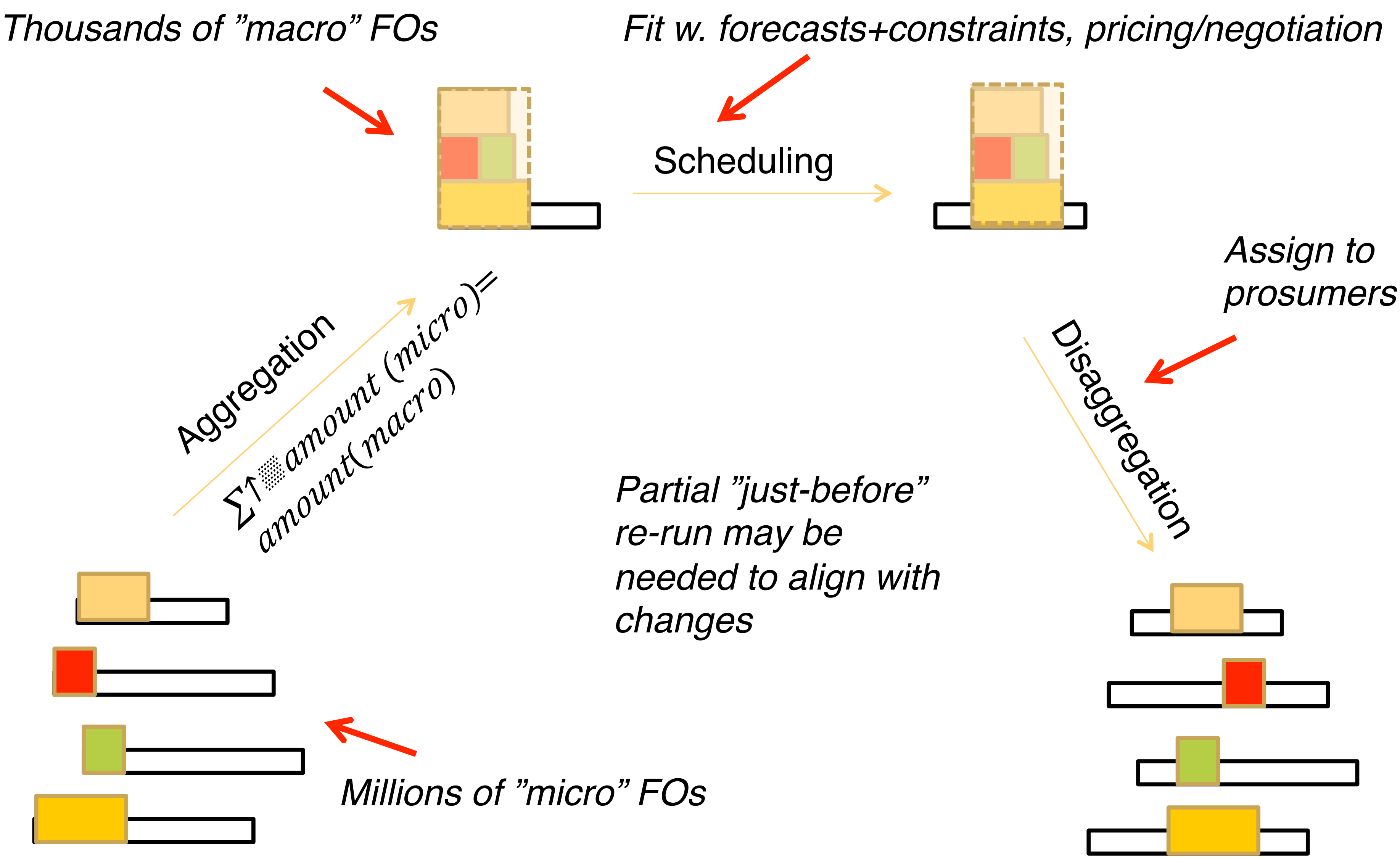


Flex-Offers: Modeling Flexibility

Standardization ongoing

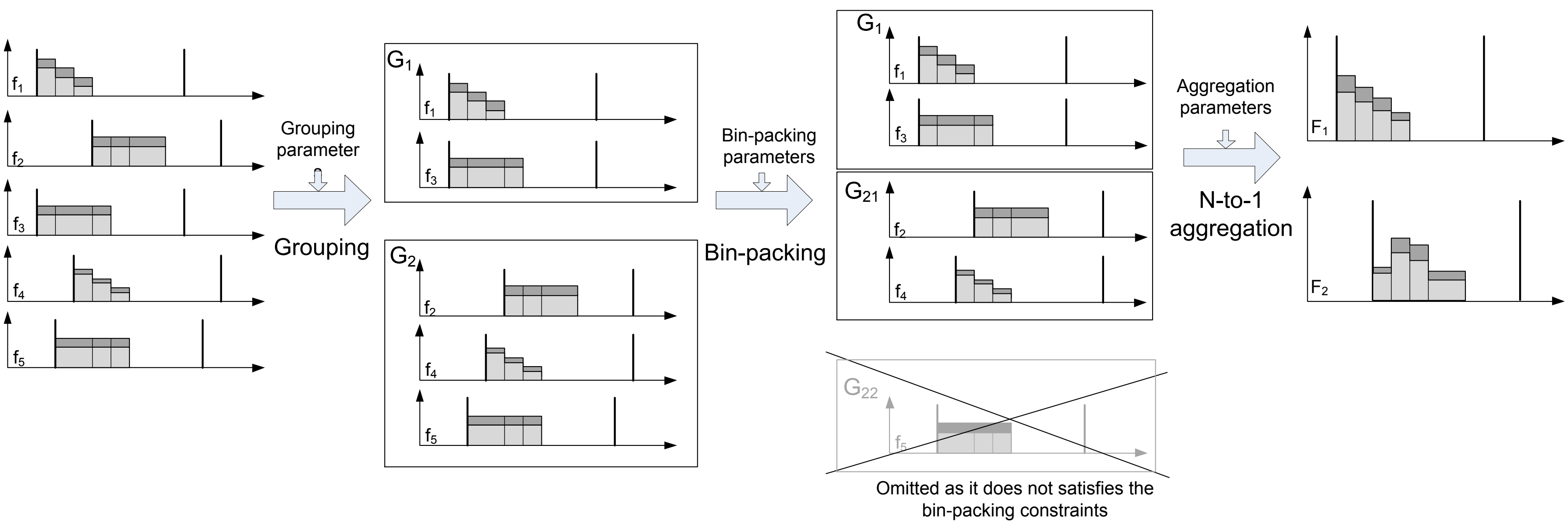


Flex-Offer Processing Cycle



Flex-offer Aggregation Overview

- Large set of input FOs aggregated into small set of output FOs
- Disaggregation does the reverse, after macro-level scheduling
- Always possible to correctly disaggregate scheduled flex-offers
- Number of aggregated flex-offers as small as possible
- Loss of flexibility in the aggregation as small as possible
- Aggregation+scheduling+disaggregation within 10 min
- 3-step aggregation (grouping, bin-packing,N-to-1 aggregation)



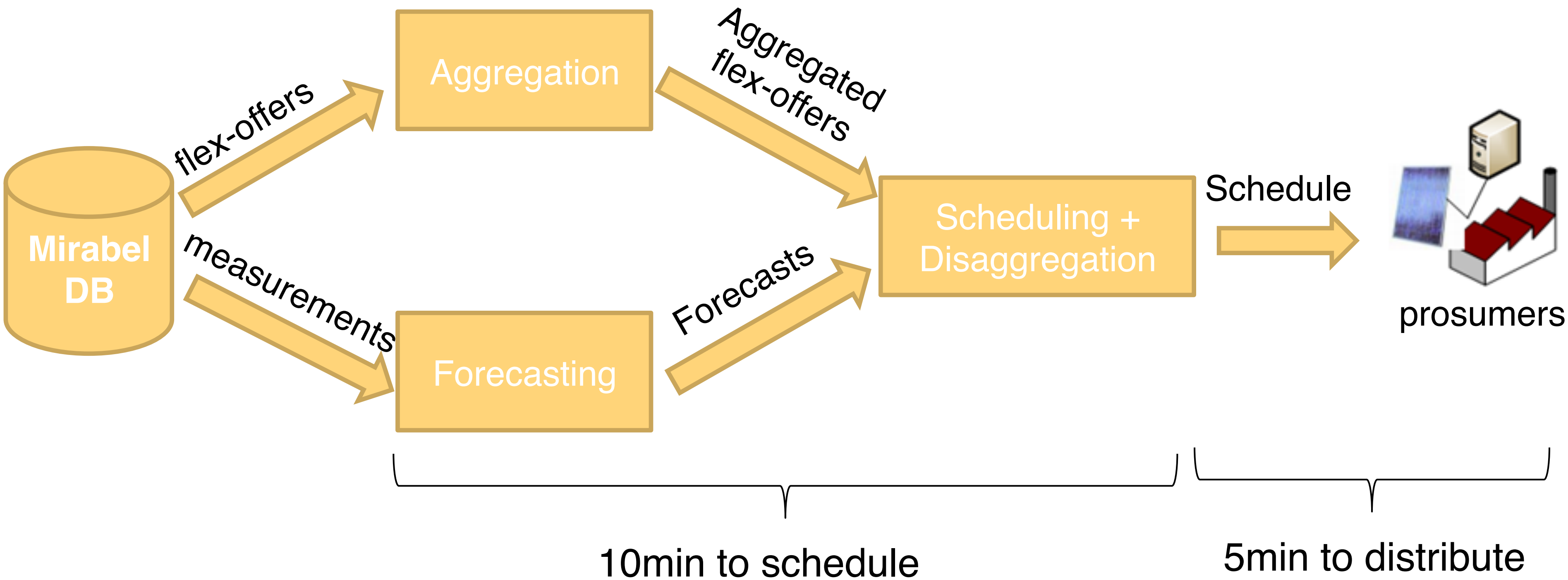
Forecasting+Scheduling Overview

- Forecasting
  - Transparent forecast model creation/usage/maintenance
  - Support for single and multi-equation models
  - Awareness for external influences, e.g., weather
  - Forecasting for demand, supply, FOs
  - Continuous evaluation and maintenance required
- Scheduling
  - Find best schedule for (agg) FOs, fix start times and energy flex.
  - Forecasted energy production, consumption, and market prices
  - Minimize the evaluation function (cost of imbalances)

$$c(S,M) = \underbrace{\sum_{\substack{i=1\\E_{\mathrm{I}}^i<0}}^{m} p_{\mathrm{I}-}^i \left|E_{\mathrm{I}}^i\right|}_{c_{\mathrm{I}-}} + \underbrace{\sum_{\substack{i=1\\E_{\mathrm{I}}^i>0}}^{m} p_{\mathrm{I}+}^i E_{\mathrm{I}}^i}_{c_{\mathrm{I}+}} + \underbrace{\sum_{k=1}^{n}\left(\sum_{j=1}^{n_k} p_k^j E_k^j\right)}_{c_{\mathrm{FO}}} + \underbrace{\sum_{\substack{i=1\\E_{\mathrm{M}}^i<0}}^{m} p_{\mathrm{M}-}^i \left|E_{\mathrm{M}}^i\right|}_{c_{\mathrm{M}-}} - \underbrace{\sum_{\substack{i=1\\E_{\mathrm{M}}^i>0}}^{m} p_{\mathrm{M}+}^i E_{\mathrm{M}}^i}_{c_{\mathrm{M}+}}$$

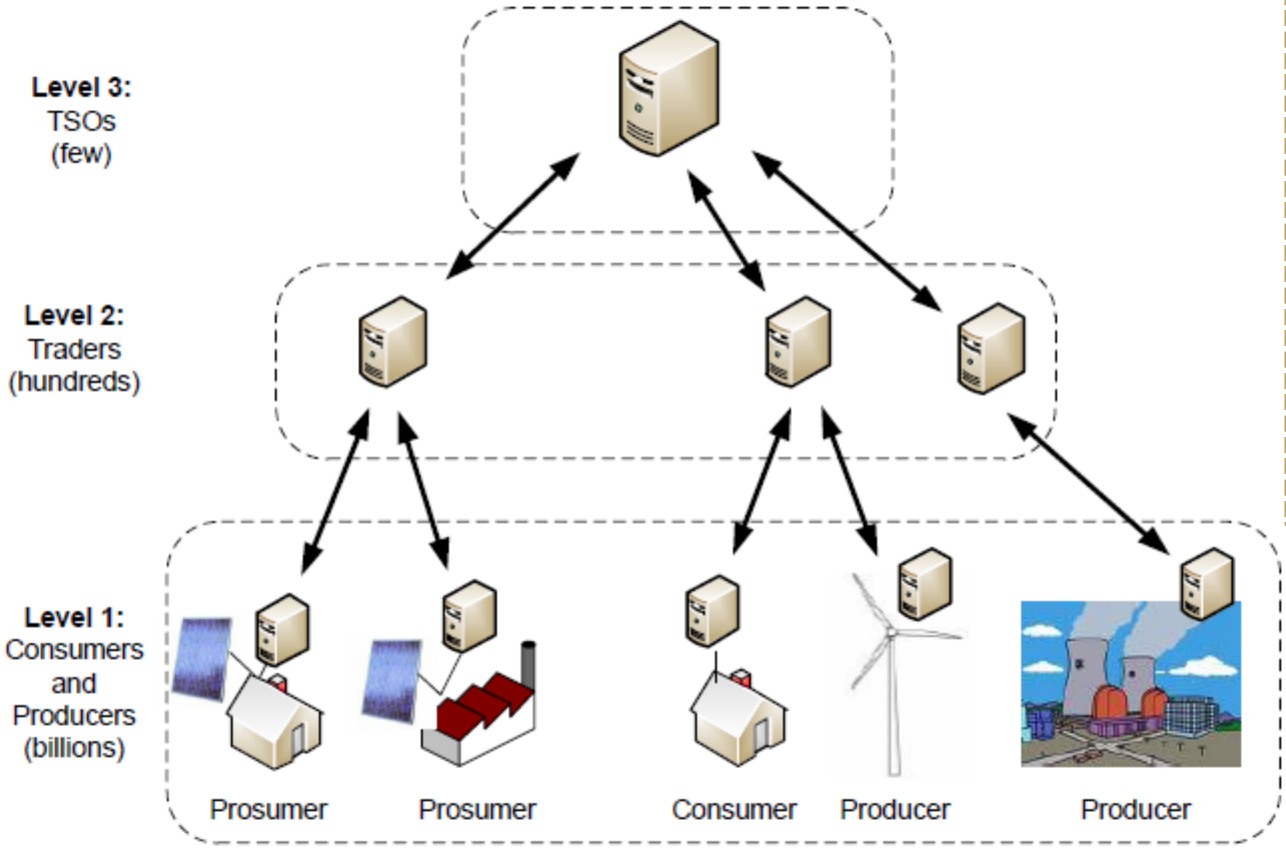
  - Prohibitive to find optimal solution, so approximation used.

Component Interplay and Timing



MIRABEL Distributed System

- Reflect the Harmonized Role Model for energy markets



Denmark

1 TSO

~100 traders

$\sim 3 \cdot 10^6$ households

⟷ Exchange of

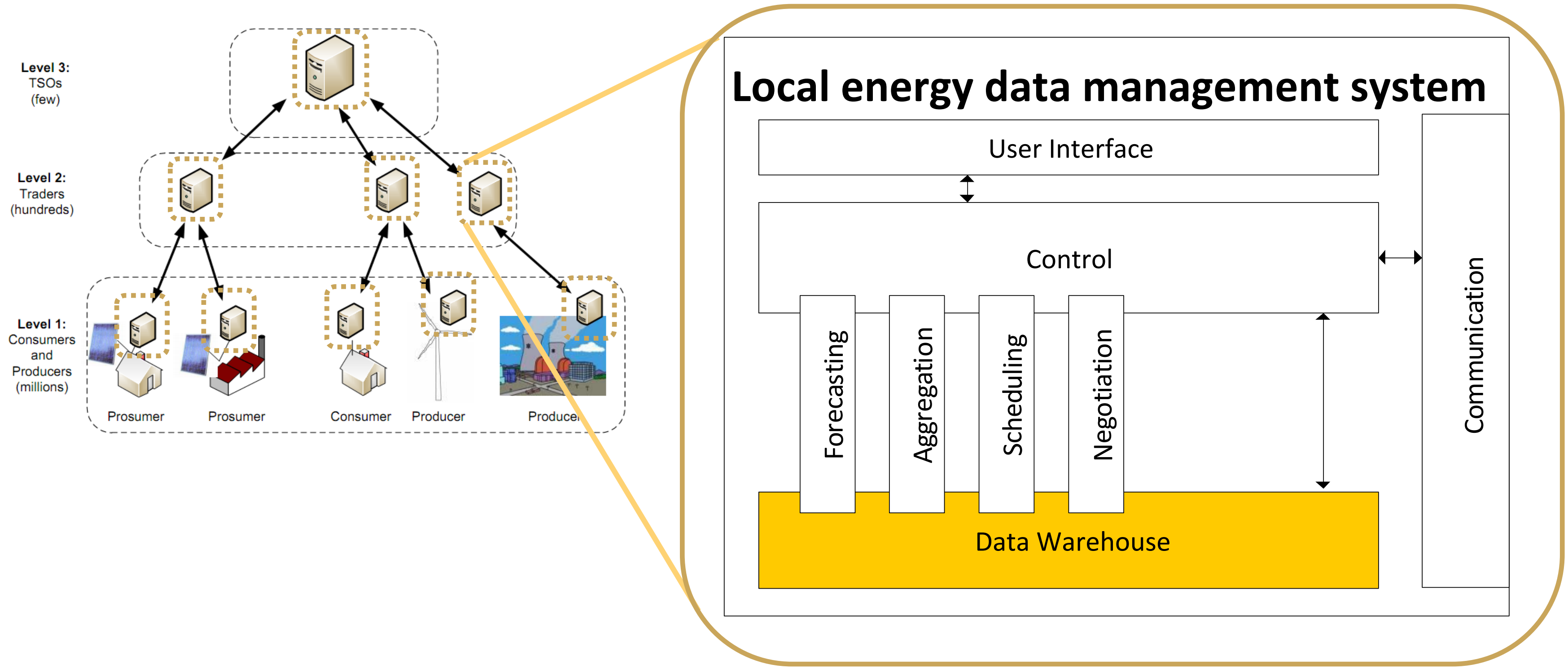
- Measurements
- Flex-offers
- Prices
- …

Flex-Offer Storage and Querying

- How to store and query flex-offers and other MIRABEL data in an object-relational data warehouse ?

MIRABEL DW Context

EDMS NODE ARCHITECTURE



- DW accepts many insert/analytical queries from analytical components
- A suitable DW schema is need for efficient query evaluation

Storage Contributions

We:

- Present a generic DW schema supporting all levels of the EDMS hierarchy
- Discuss the complexities of the schema compared to traditional DW schemas
- Discuss alternative data modeling strategies
- Evaluate schema alternatives using typical queries from the MIRABEL project

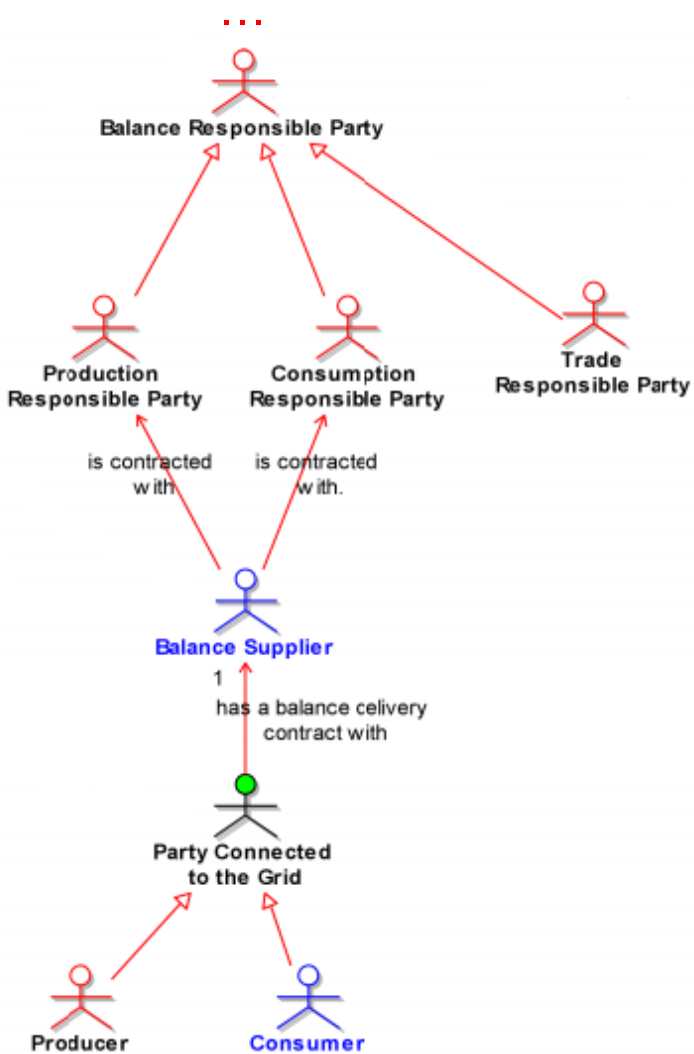
- More on the MIRABEL EDMS: “Data Management in the MIRABEL Smart Grid System”, EnDM 2012
- More on the MIRABEL DW: “Real-time Business Intelligence in the MIRABEL Smart Grid System”, BIRTE 2012

MIRABEL DW: Schema

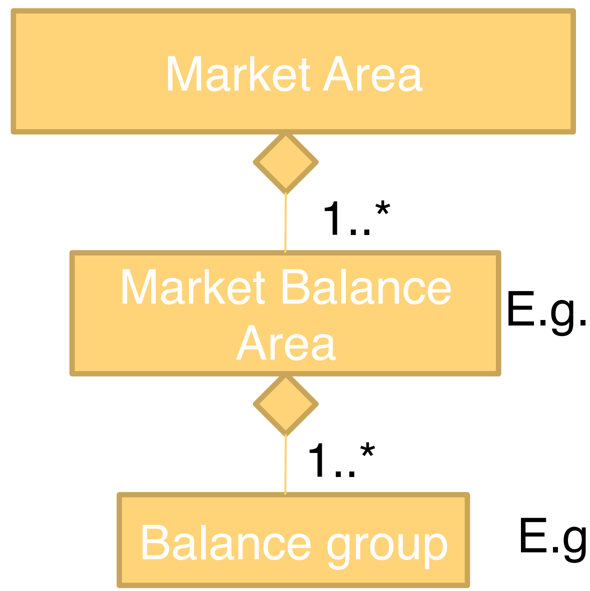
- Based on the MIRABEL data model
  - Common information model (CIM) by IEC
    - Represent major objects in an electric utility enterprise
  - Harmonized Electricity Market Role Model by ebIX®, EFET and ENTSO-E
    - Define administrative data internally interchanged between European electricity markets

  None of the existing models focus on storage of energy-related entities
- Schema is complete for the prototype of the MIRABEL system
- Represents energy data essential in the MIRABEL context
  - Actors of European Electricity Market,
  - Flex-offers,
  - Time series of energy, power, and prices

MIRABEL DW: Actors and Roles



- There are many roles
- A role:
  - may have specializations
  - may interact with other roles
  - belongs to one of the logical areas



E.g, Nordpool

E.g., Eastern Denmark

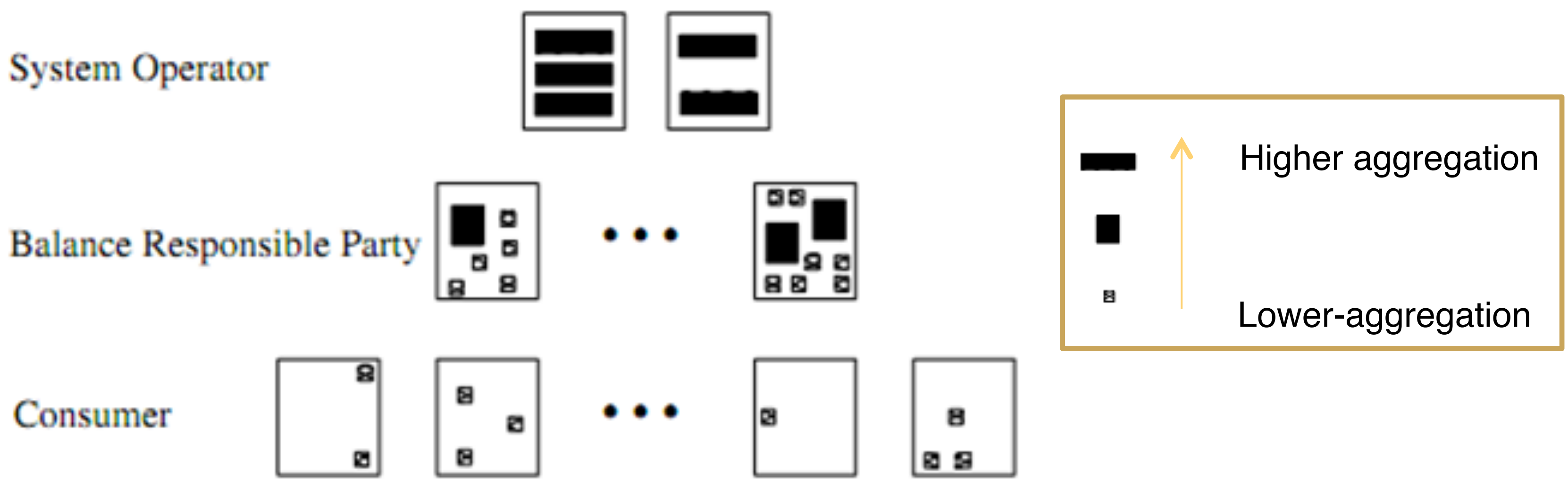
E.g., Vattenfall BG

- An actor plays one or more roles

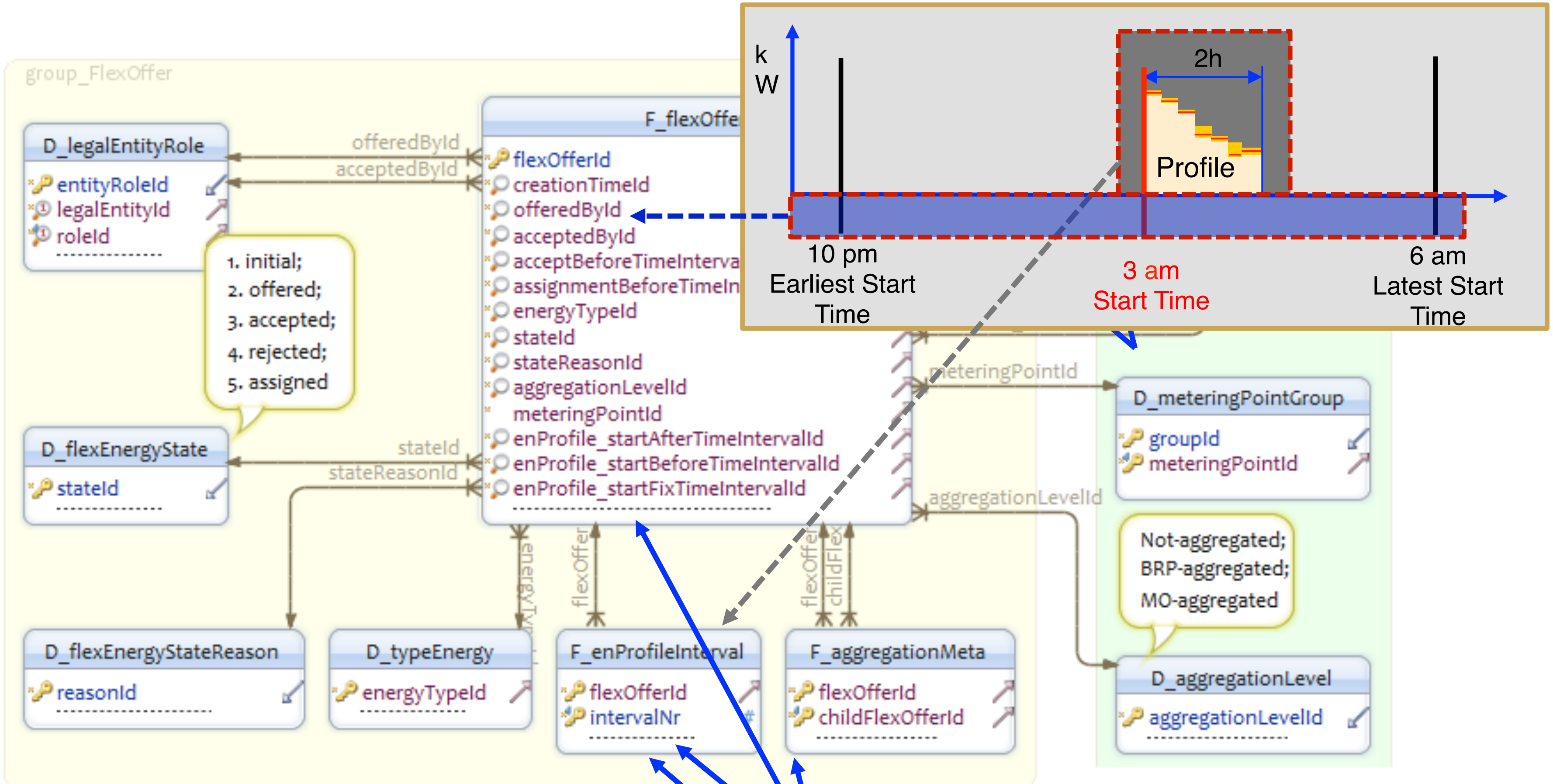
We model a market area with actors + roles

MIRABEL DW: Actors and Roles

- For every actor-role, the schema captures:
  - Time-series
  - Flex-offers



MIRABEL DW: Flex-Offers



Properties:

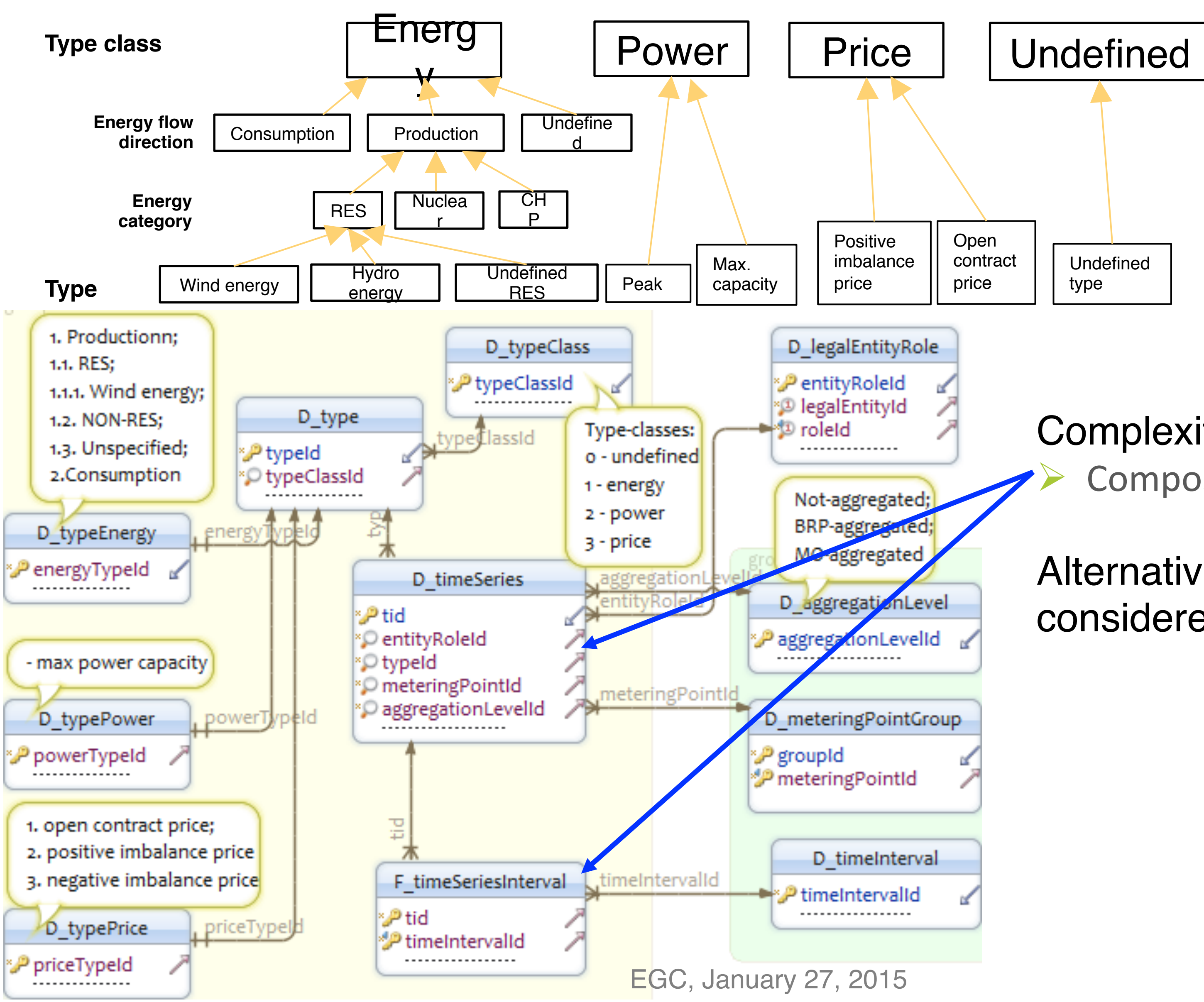
- Time is discretized
- Flex-offers of different aggregation levels
- Instances of flex-offers represented

Complexities:

- Non-atomic composed facts
- Facts about facts

Alternative designs are considered

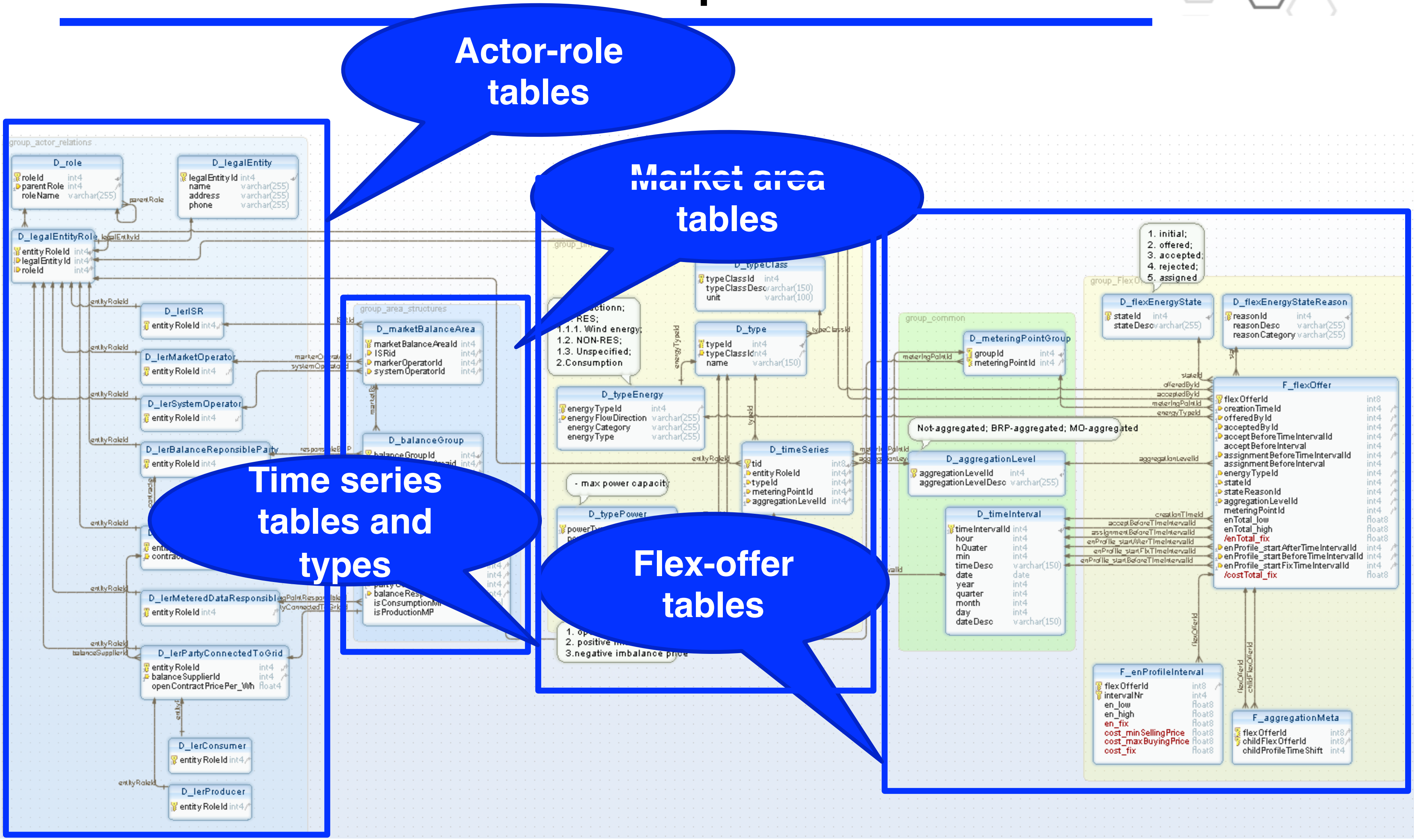
MIRABEL DW: Time Series



Complexities:
- Composed facts

Alternative designs are considered

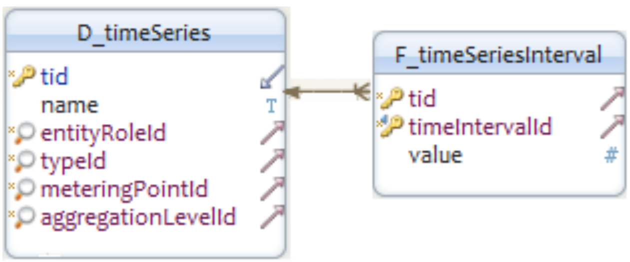
MIRABEL DW: Complete Schema

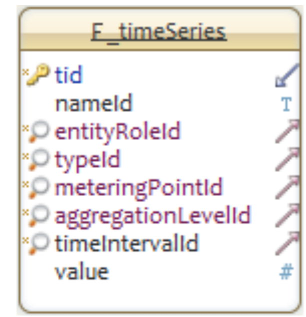


MIRABEL DW: Alternative Designs

Flex-offer and timeseries schema alternatives

– Denormalized





D_timeSeries

tid	name	entityRoleID	typeId	…
1	TS1	0	1113	
2	TS2	1	1114	

F_timeSeriesInterval

tid	timeIntervalId	value
1	1000	11.2
1	1001	11.4
2	1000	101.1
2	1001	101.2

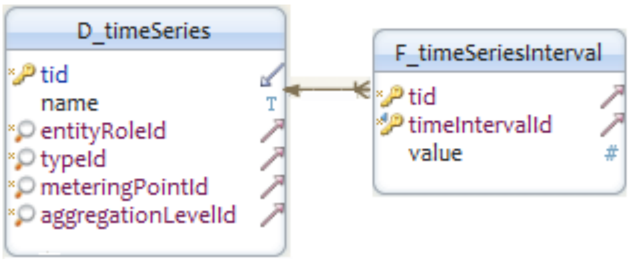
F_timeSeries

tid	name	entityRoleId	typeId	…	timeIntervalId	value
1	TS1	0	1113		1000	11.2
1	TS1	0	1113		1001	11.4
2	TS2	1	1114		1000	101.1
2	TS2	1	1114		1001	101.2

MIRABEL DW: Alternative Designs

Flex-offer and timeseries schema alternatives

– Array-based

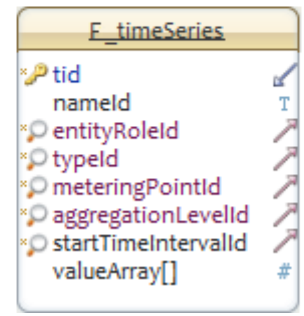


D_timeSeries

tid	name	entityRoleID	typeId	…
1	TS1	0	1113	
2	TS2	1	1114	

F_timeSeriesInterval

tid	timeIntervalId	value
1	1000	11.2
1	1001	11.4
2	1000	101.1
2	1001	101.2



F_timeSeries

tid	name	entityRoleId	typeId	…	startTimeIntervalId	valueArray
1	TS1	0	1113		1000	{ 11.2, 11.4 }
2	TS2	1	1114		1000	{ 101.1, 101.2 }

MIRABEL DW: Experiments

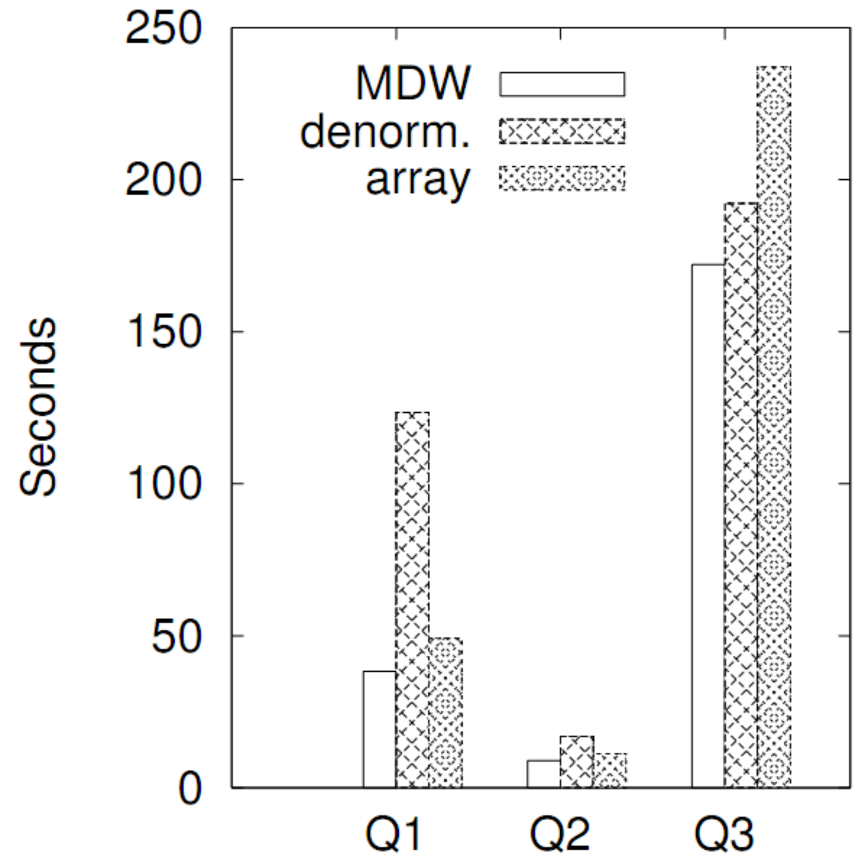
Experiment setup

- Real energy consumption data: 963 time series, 32.1M values (MeRegio),
- Synthetically generated 3.1M flex-offers
- Standard server machine
  - Linux server with 16 GB RAM, 2x Intel Xeon CPUs, 4 SATA 7200RPM disks
  - PostgreSQL 9.1, tables are “fully vacuumed”
- Queries executed in round-robin fashion 5 times

Flex-Offer Schema Experiments

Flex-offer queries

- Q1: Compute total flexibility per flex-offer
- Q2: Compute sum of all scheduled (fixed) energy
- Q3: Builds a time series that represents amounts of scheduled (fixed) energy



Results

- MDW variant is the fastest
- MDW variant uses optimal amount of space

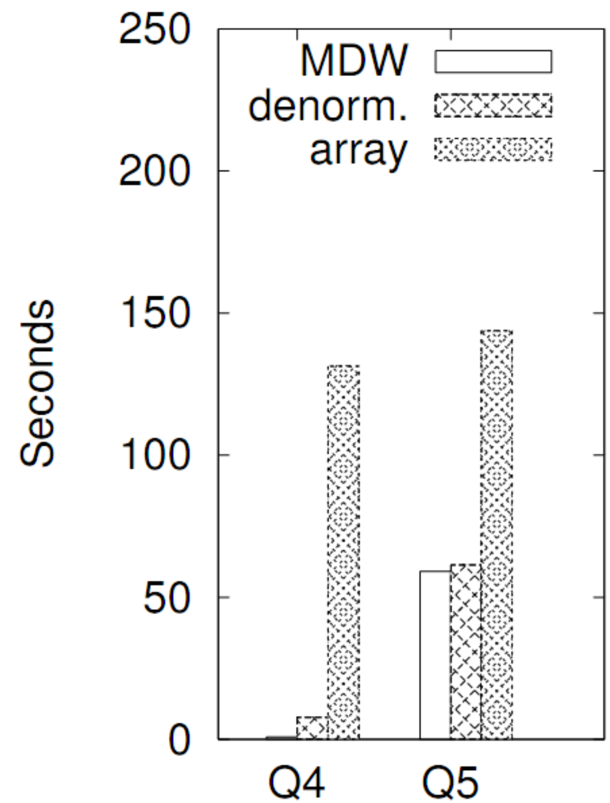
Time Series Schema Experiments

Time series queries

- Q4:Compute energy balance for 24h considering total demand and supply
- Q5: Find time series with average energy exceeding an average time series by 25%

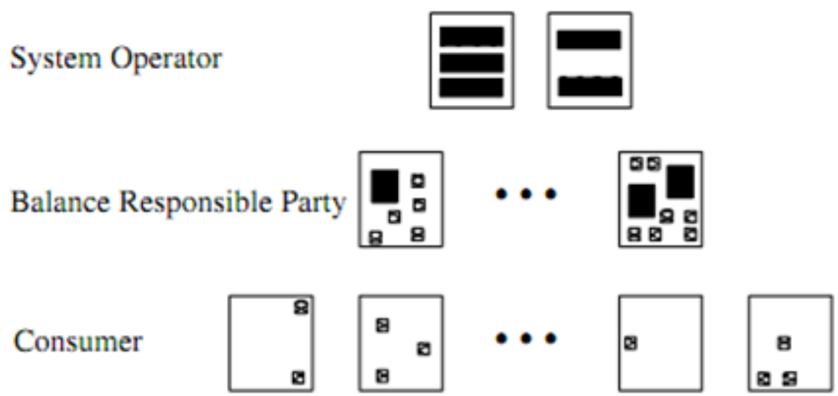
Results

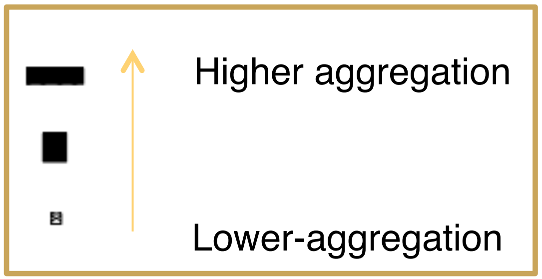
- MDW variant is the fastest
- MDW variant uses optimal amount of space



MIRABEL DW: Research Directions

- (Future) distribution of DW
  - The schema will be replicated on all nodes of EDMS
  - Node holds only relevant data and of specific granularity





- Challenges
  - Propagation of data through the hierarchy, caching
  - Specialized versions of the schema for different types of nodes such that queries formulated on generic schema can be translated to the specialized schemas

MIRABEL DW: Conclusions

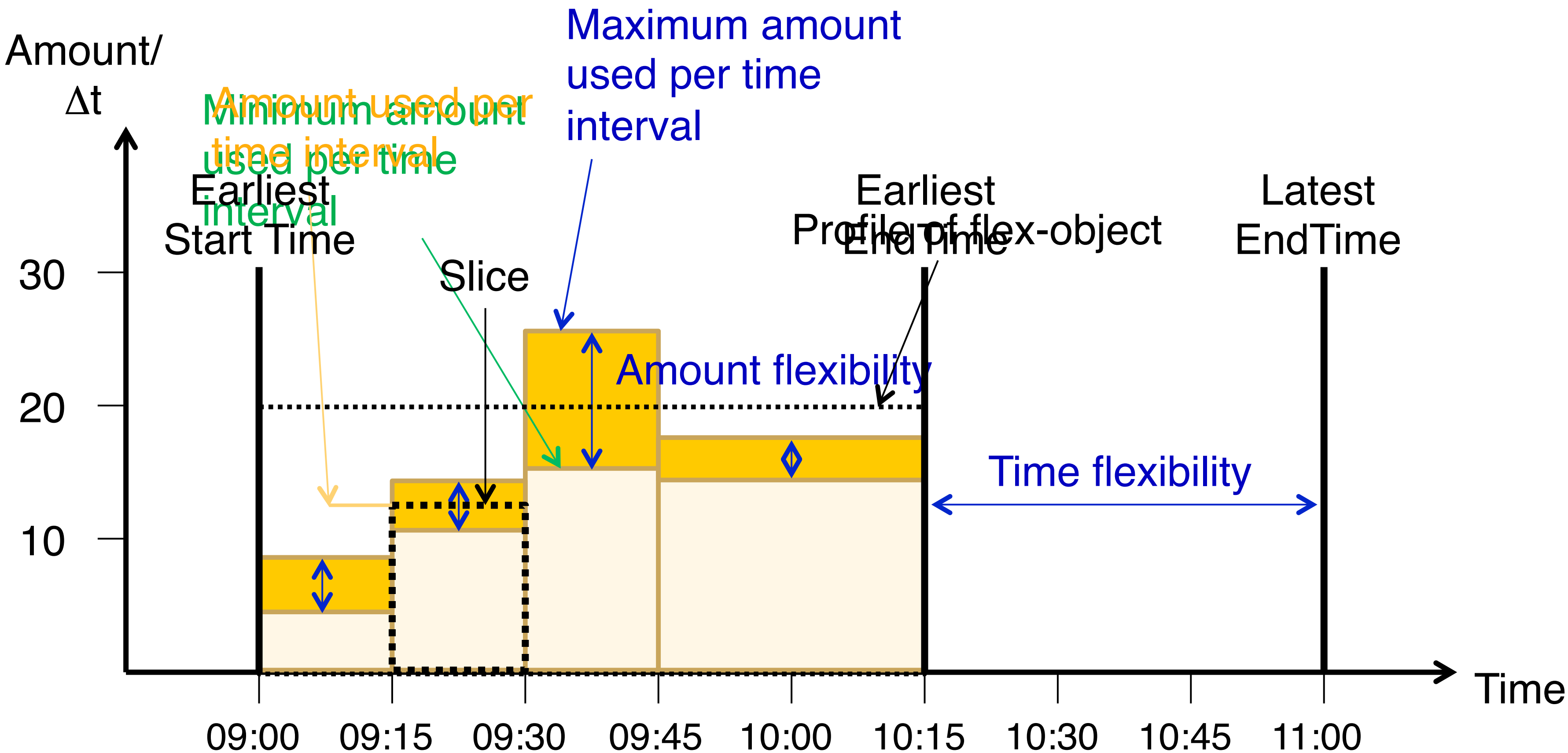
- Designed a generic DW schema for complex energy data
- The schema has a number of interesting complexities
  - Facts about facts
  - Composed non-atomic facts
- The schema can be used by a different nodes of hierarchical system
- Evaluated different alternatives (denormalization, arrays)

Aggregating Flex-Offers

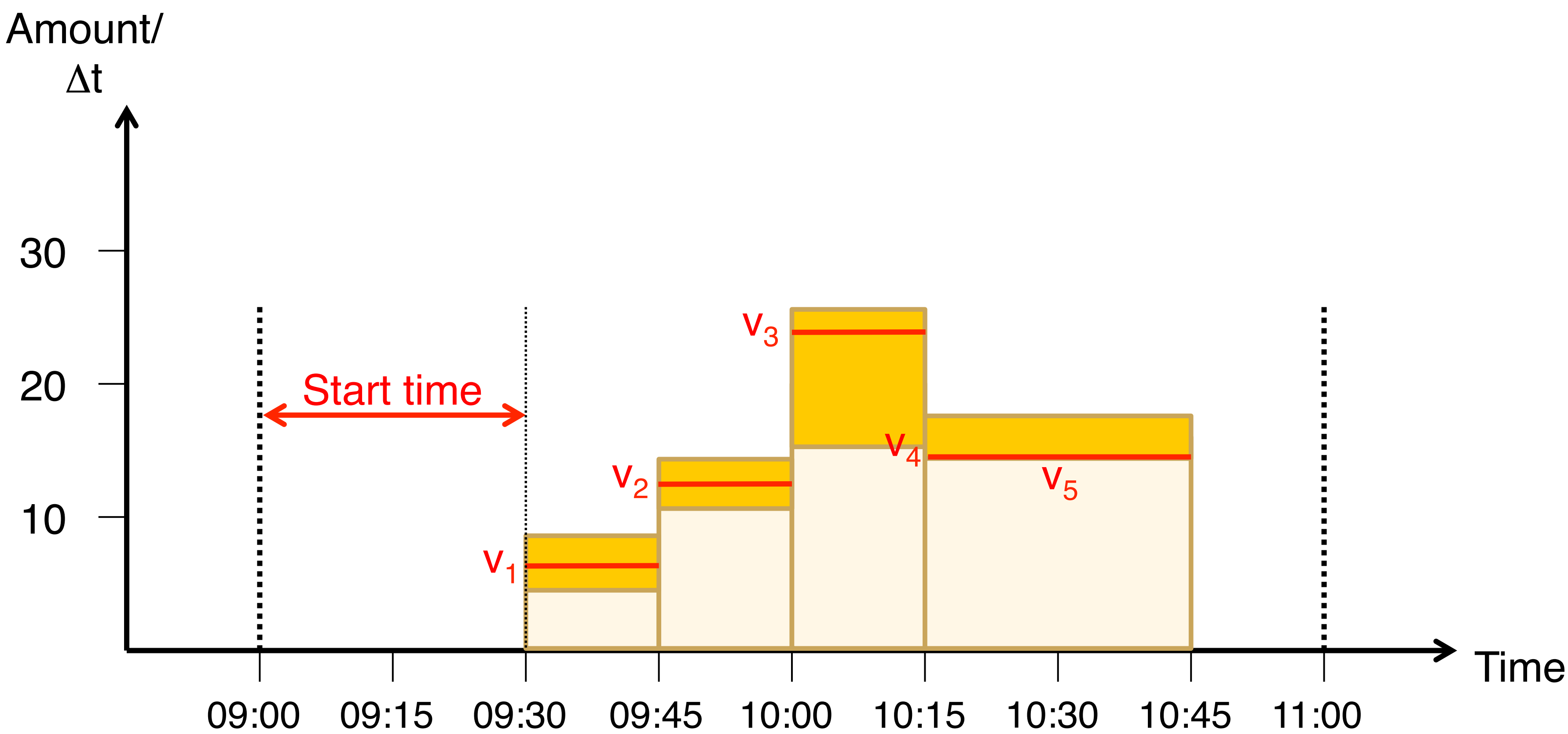
- How to we aggregate and disaggregate flex-offers?
- How do we compose many small units of flexibility into fewer, larger, and more useful units, while retaining most of the flexibility ?

Flex-Object (Generalization)

- Flexibility object (flex-object) represents the usage of a resource (e.g., energy) over time as well as flexibilities



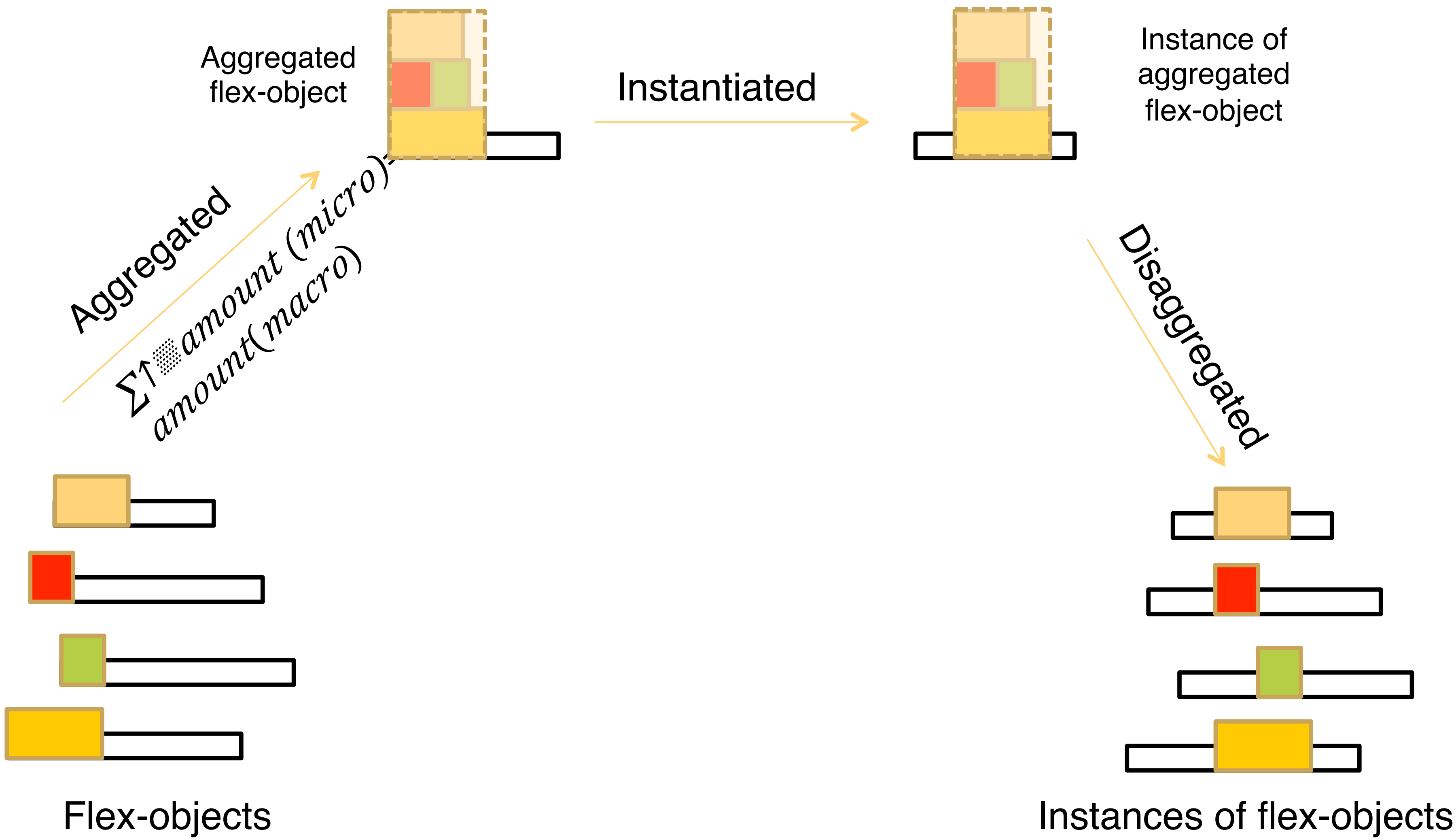
Flex-Object Instance



Flex-Object Database Vision

- The energy management system of the utility company manages a large number of flex-objects
- Flex-object database is needed:
  - Flex-objects as first-class citizens
  - Dedicated or storing other types of data
  - Supported functionality:
    - Different types of flexibility
    - Complex hierachies such as energy distribution grids
  - Supported queries:
    - Flexibility availability queries – *min/max amounts available at a time interval*
    - Adjustment potential queries - *distribution of amounts that can be potentially injected into (or extracted from) a given time interval*
    - Fixing queries - *alter the plan based on the amount to inject or extract*
    - Scheduling queries – *instantiates flex-objects to match a time series*
    - <u>Flex-object aggregation queries</u> – *combines "micro" flex-objects into fewer "macro" flex-objects*
    - <u>Flex-object disaggregation queries</u> - explode an instance of a "macro" flex-object into instances of "micro" flex-offers

Flex-offer (FO) life cycle Recap



Problem definition

Aggregation

- Takes N and produces M flex-objects

Disaggregation

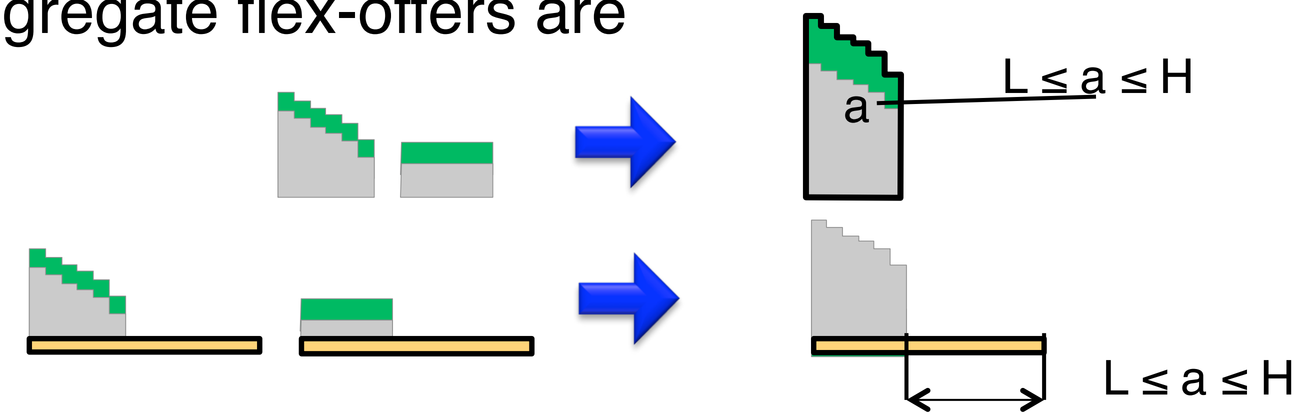
- Takes M and produces N instances of flex-objects

$$\boxed{M \ll N}$$

Problem definition

Additional requirements for aggregation and disaggregation:

- Compression and flexibility trade-off requirement
- Aggregate constraint requirement, e.g., to limit "how big" aggregate flex-offers are



- Incremental Update Requirement (for the online scenario)
  - New flex-objects are continuously received
  - *Earliest starting time* of existing flex-objects are approaching

Need to be able to efficiently integrate changed flex-objects into aggregates

Aggregation solutions

Three solutions presented in the paper:

- **N-to-1 aggregation**
  - **+ Satisfies the amount balance requirement**
  - – Does not satisfy compression/flexibility, aggregate constraint, and the incremental updates requirement
  - – Loses most of flex-object flexibility

- **N-to-M aggregation**

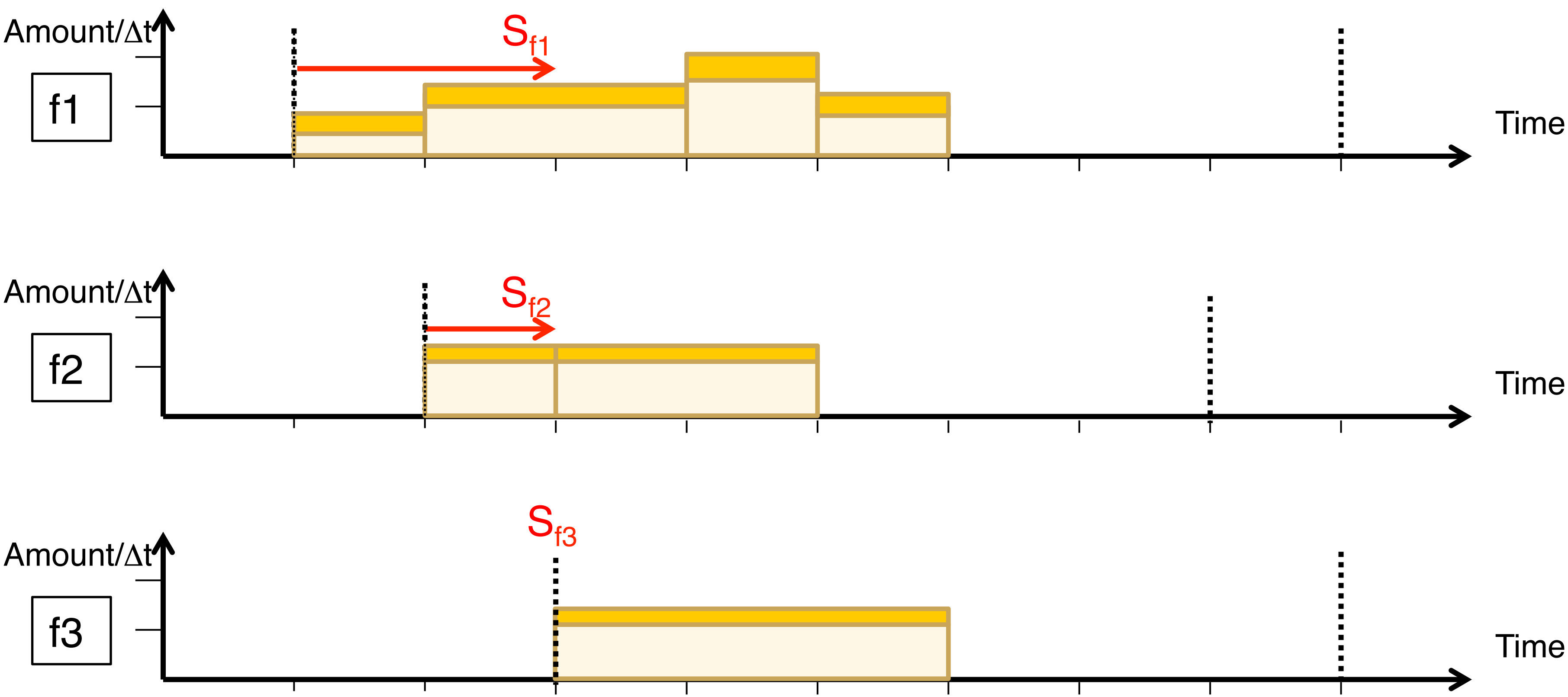
  Based on prior grouping and bin-packing
  - **+ Satisfies compression/flexibility, aggregate constraint requirements**
  - – Does not satisfy the incremental update requirement

- **Incremental N-to-M aggregation**
  - **+ Satisfies all requirements**

N-to-1 aggregation

N-to-1 aggregation

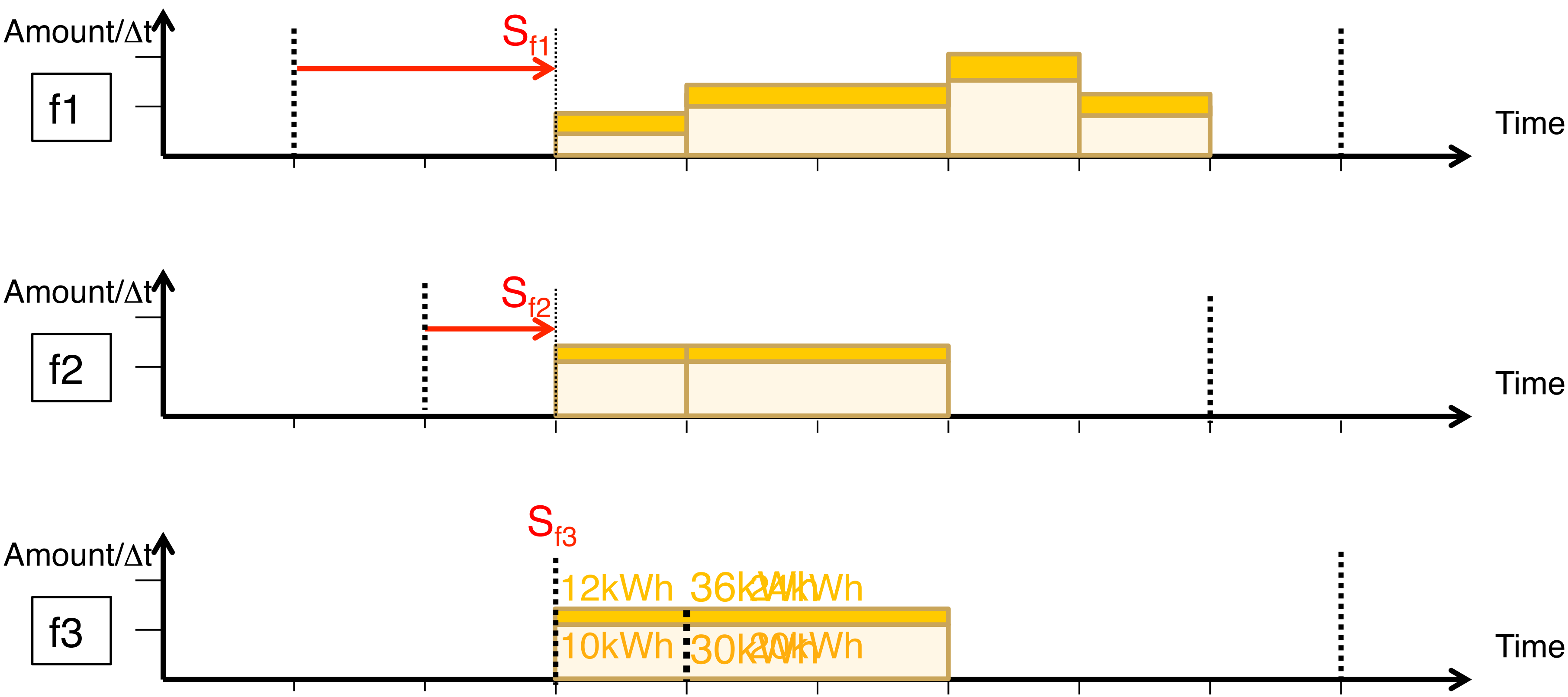
- To aggregate flex-objects, we follow these steps
  - Align profiles (partially instantiate flex-objects)



N-to-1 aggregation

N-to-1 aggregation

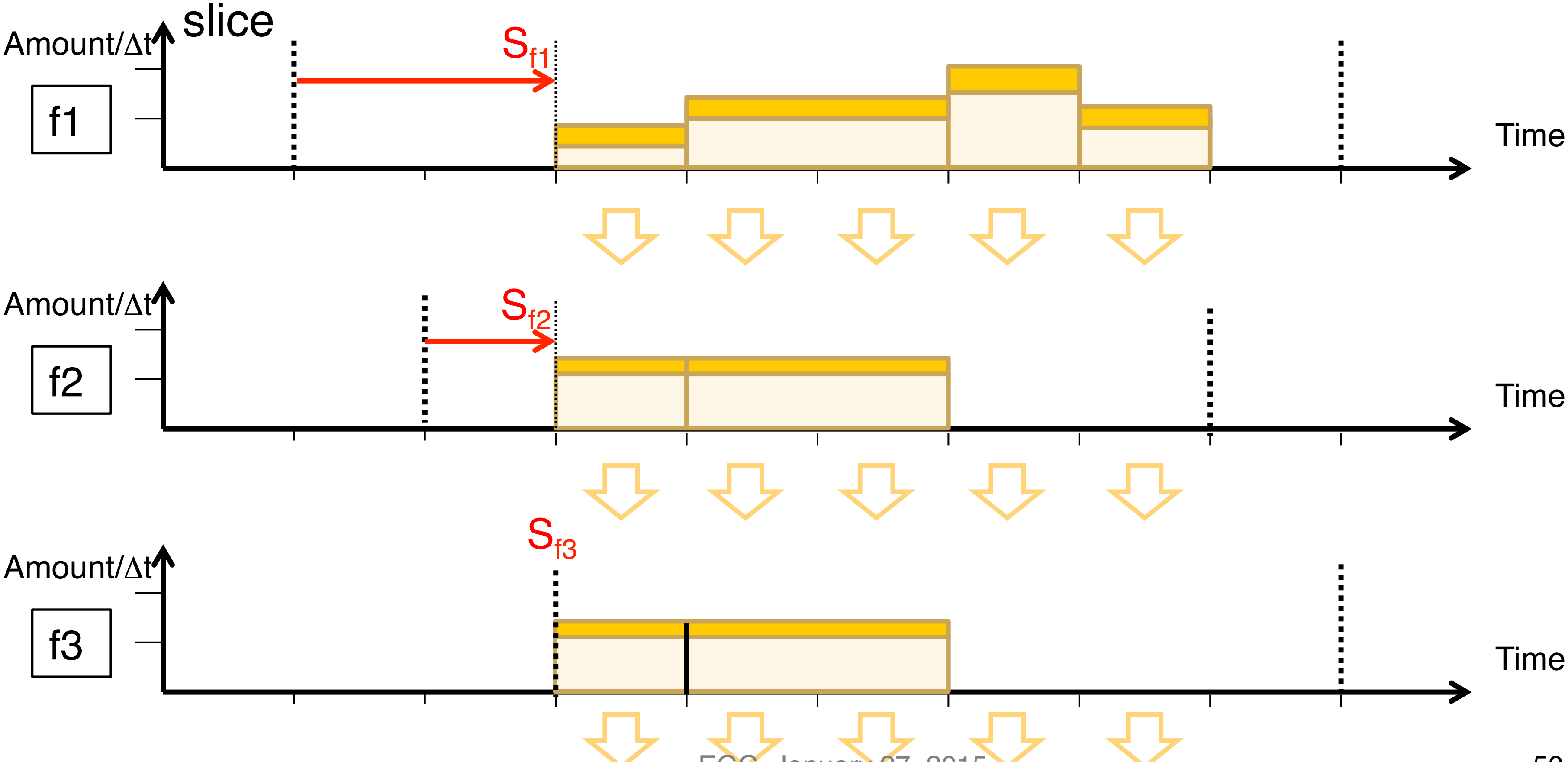
- To aggregate flex-objects, we follow these steps
  - Partition slices if needed



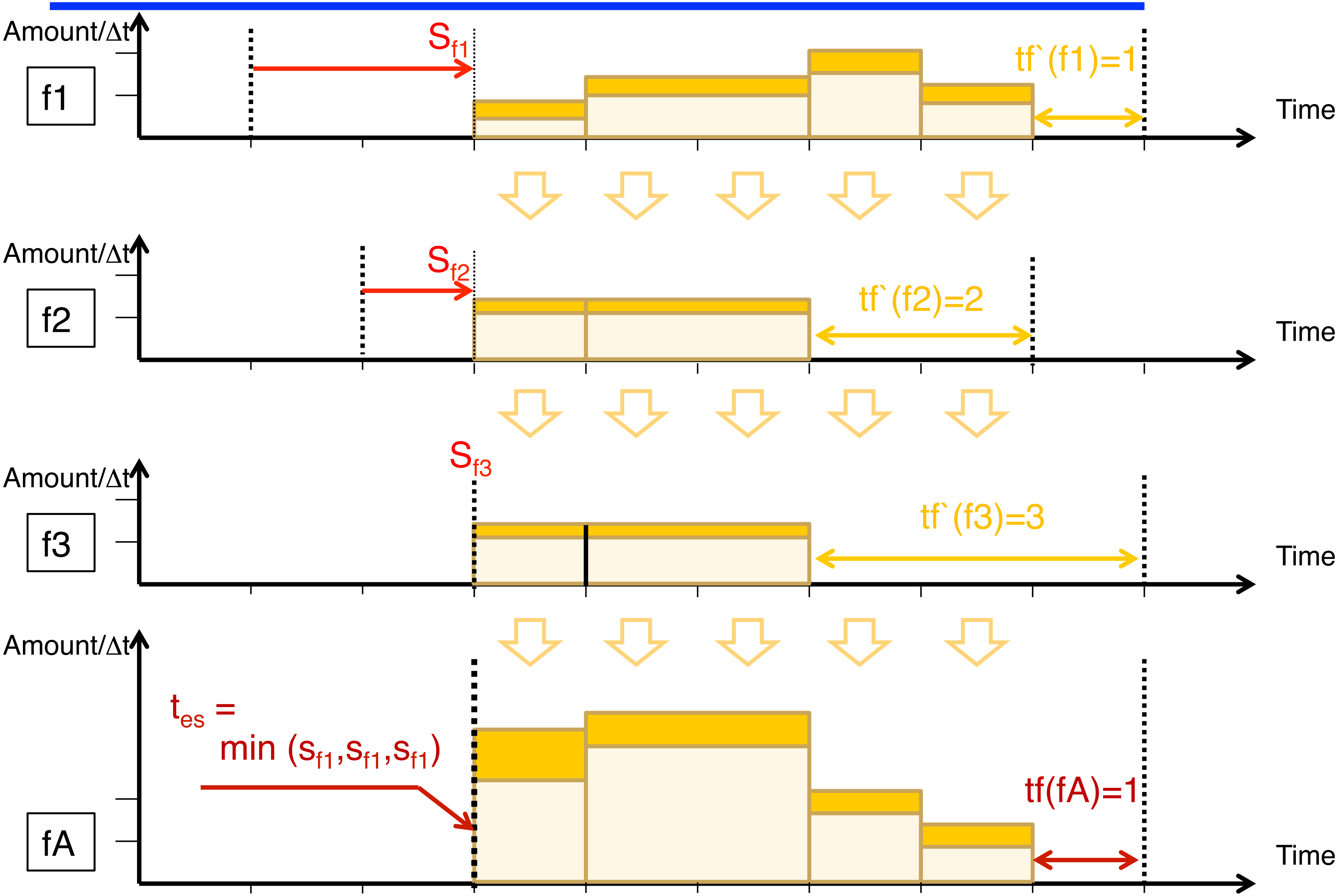
N-to-1 aggregation

N-to-1 aggregation

- To aggregate flex-objects, we follow these steps
  - Build a new profile by summing all corresponding amounts for each slice



N-to-1 aggregation



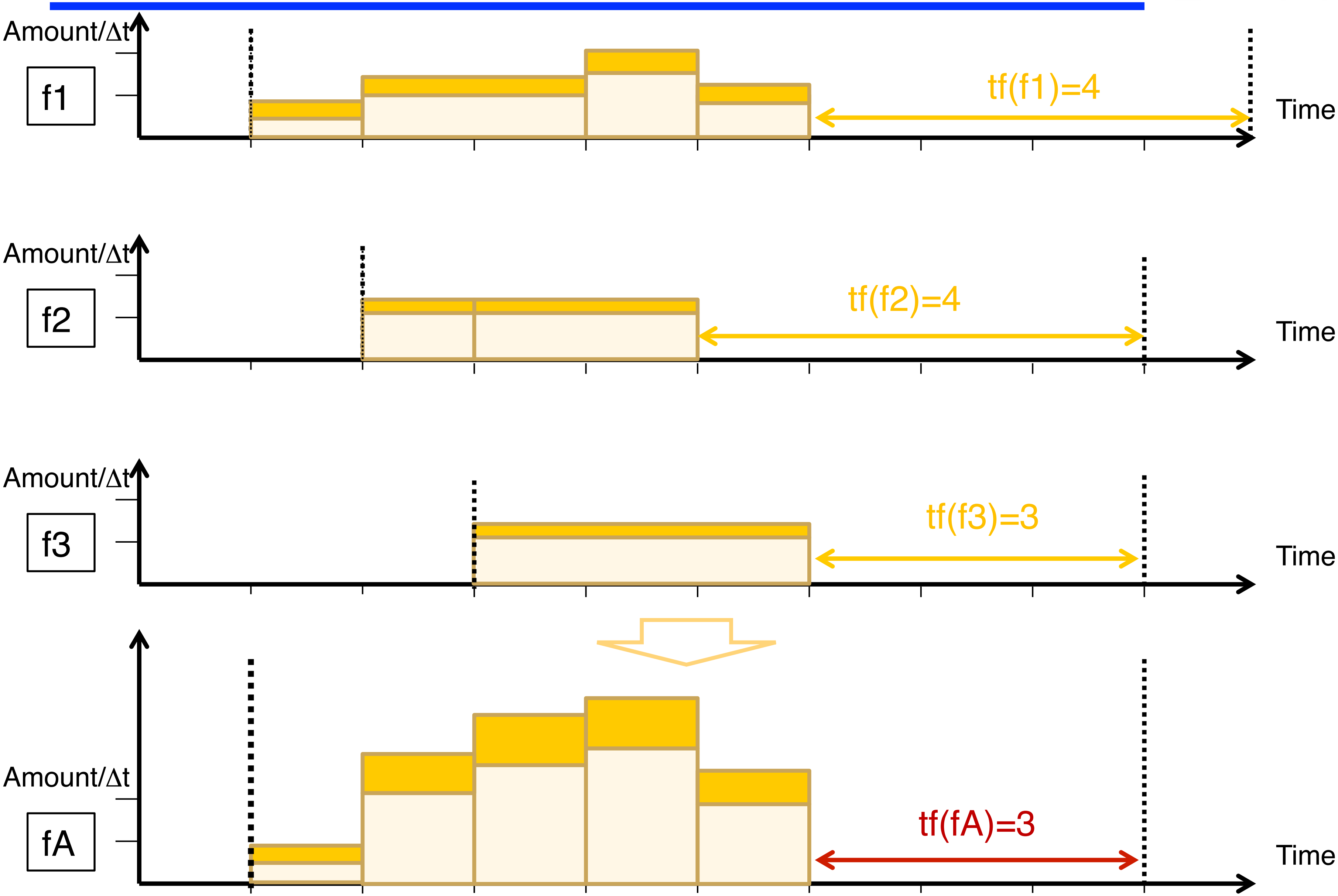
N-to-1 aggregation

- Different alignments result in different shapes of profiles and remaining time flexibilities, e.g.,
  - As in previous example, tf(f1) = tf(f2) = tf(f3)=3, but tf(fA)=1.
- The idea is to allow alignment such that

$$Tf(fA)=\min_{(f \text{ in } F)}\{tf(f)\}$$

- Three most important alignments ensuring this property:
  - Start-alignment
  - Soft left-alignment
  - Soft right-alignment

N-to-1 aggregation: start-alignment



N-to-1 aggregation: Start-alignment

- Start-alignment

  Pros

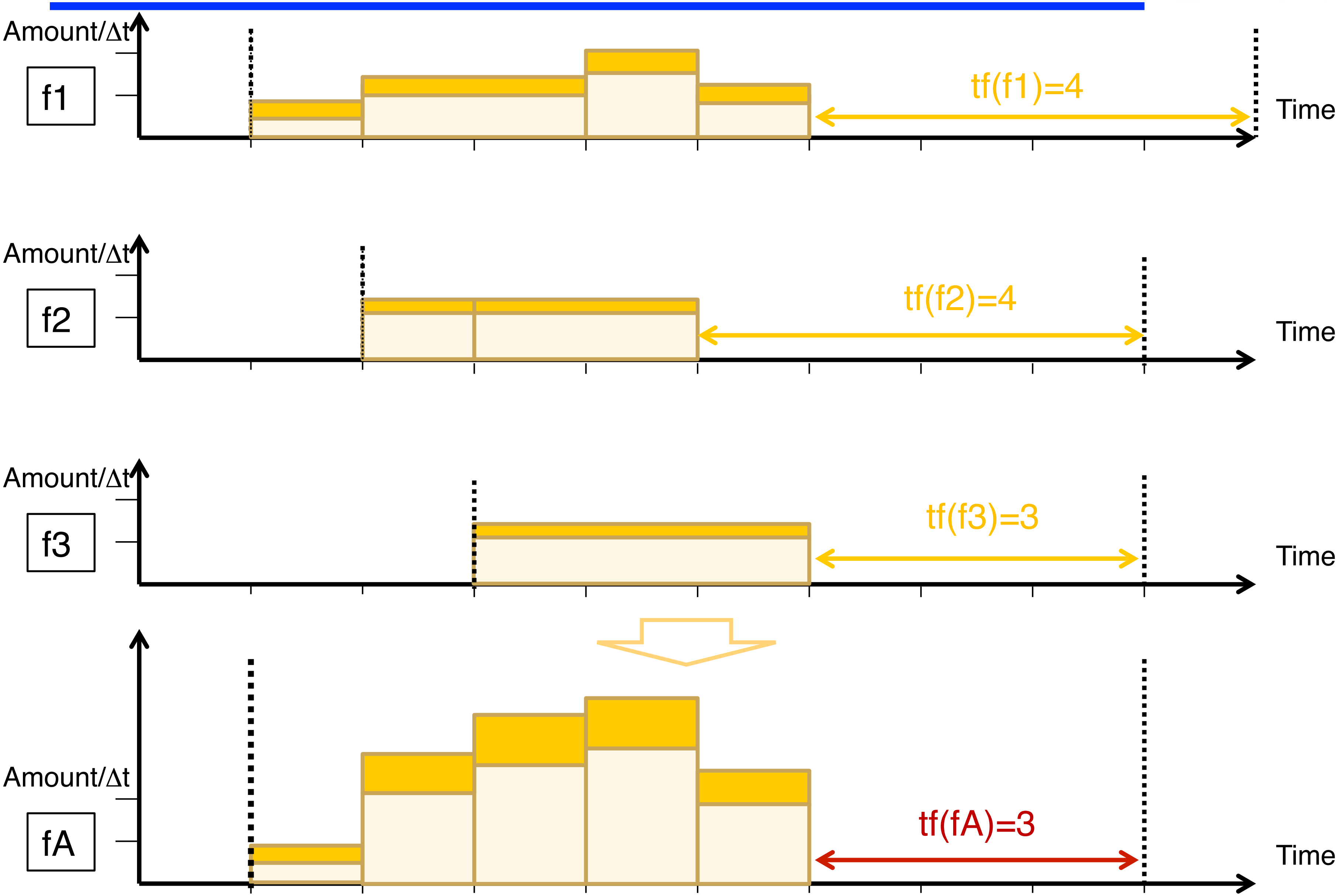
  – Spreads out amounts throughout the time extent of all individual flex-objects

  – Makes larger amounts available as early as possible

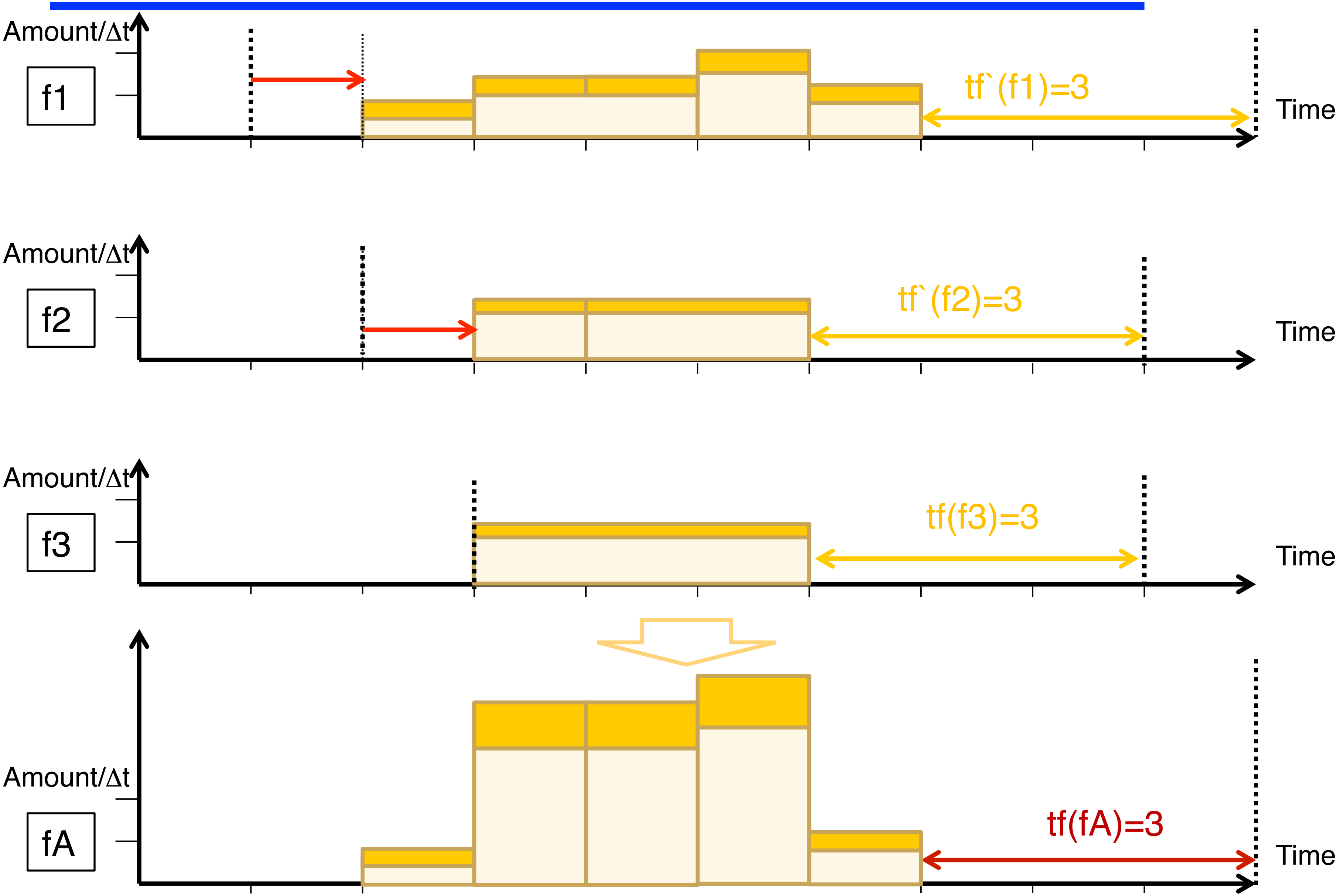
  Cons

  – Might result in very long profiles, which might be inconvenient to handle

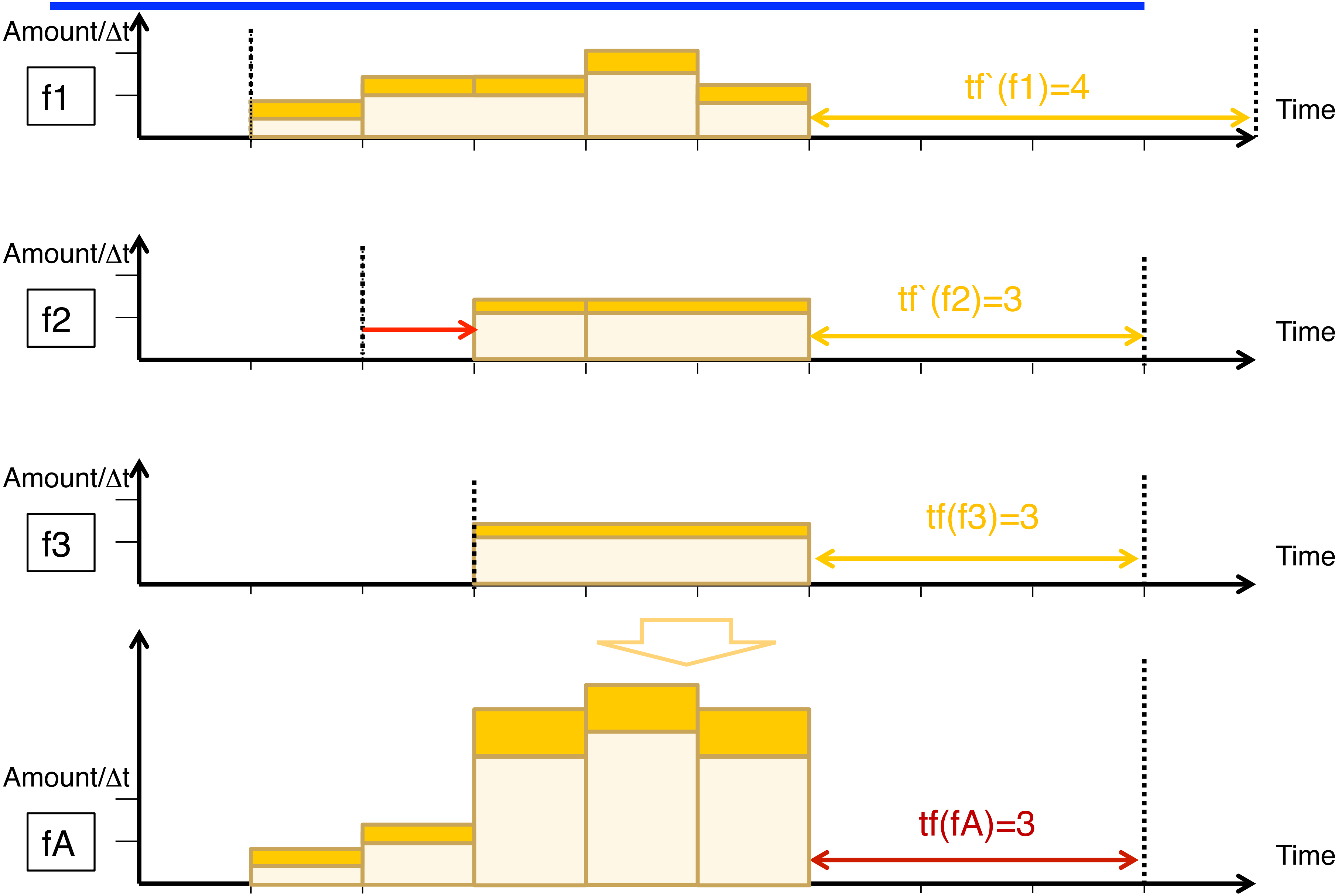
N-to-1 aggregation: soft left-alignment



N-to-1 aggregation: soft left-alignment



N-to-1 aggregation: soft right-alignment



N-to-1 aggregation

- Soft-right/left alignment

  Pros

  – Short profile with concentrated amounts in the left/right

  Cons

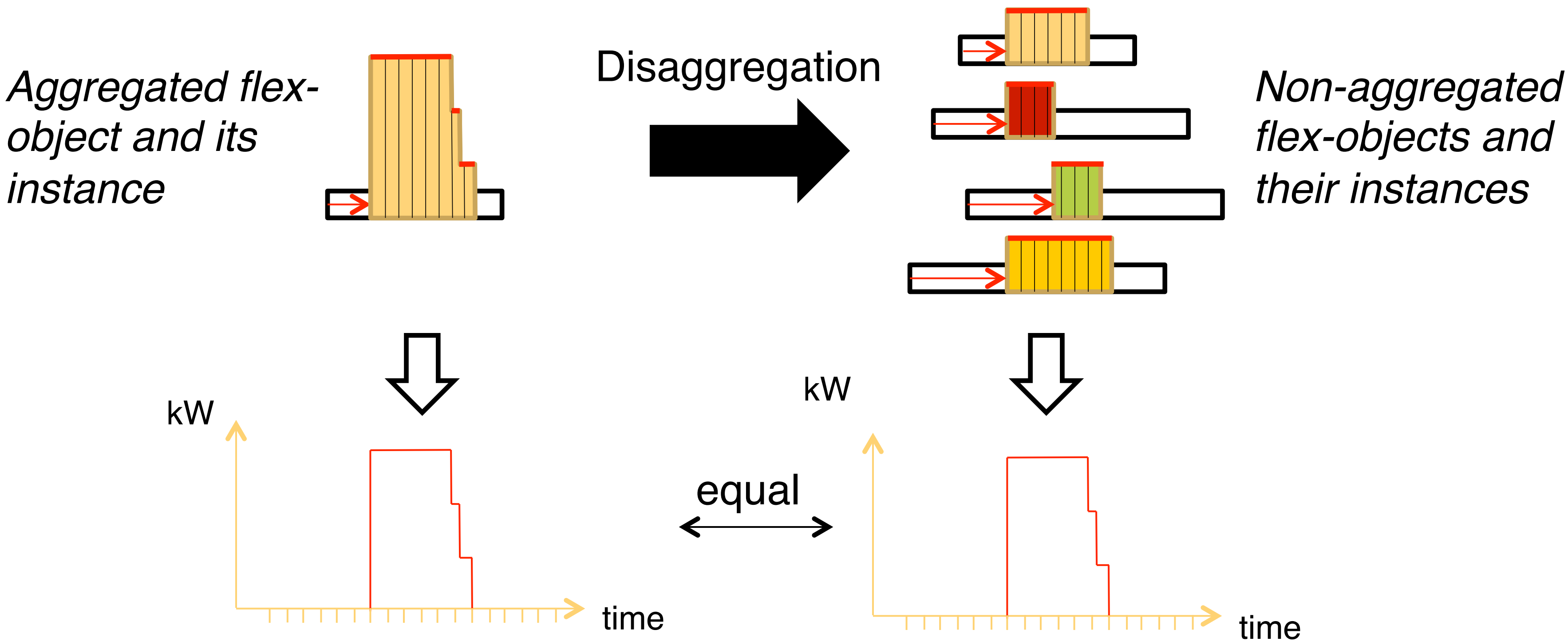
  – Not always possible to achieve hard left/right alignments

  – Amounts are not availabe early in time

- Summary of alignments
  – Time flexibility of an aggregate depends on the flex-object with smallest time flexibility

Disaggregation

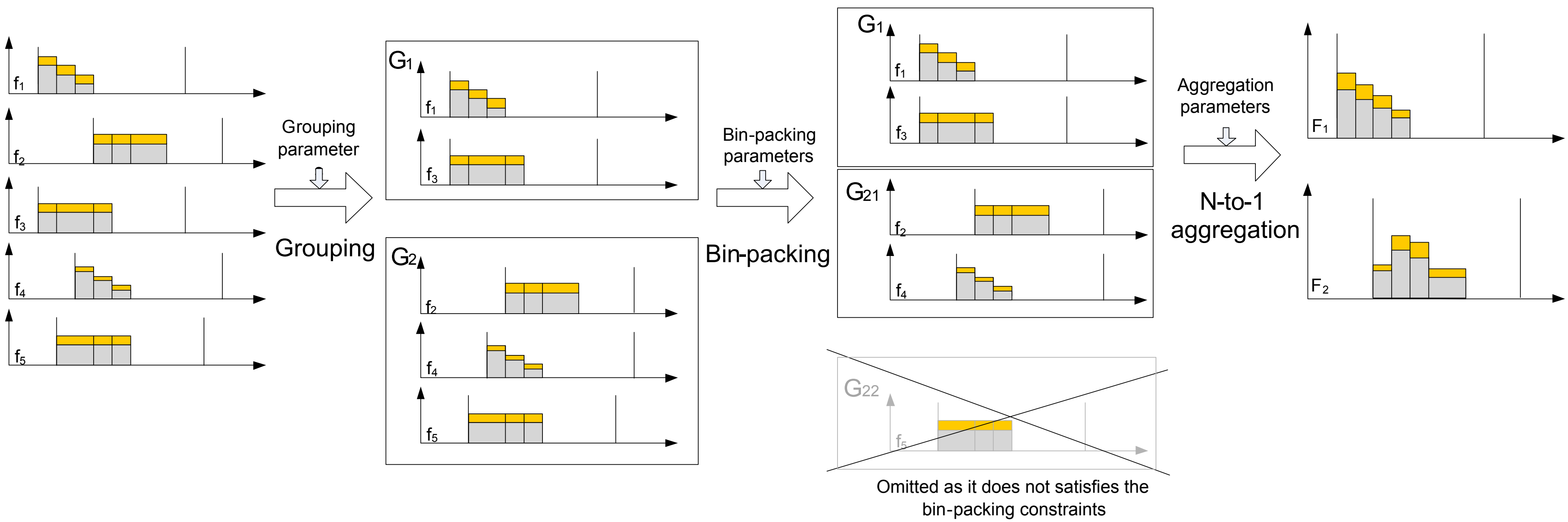
- N-to-1 aggregation is conservative
- Disaggregation is feasible for every instances of aggregated flex-objects



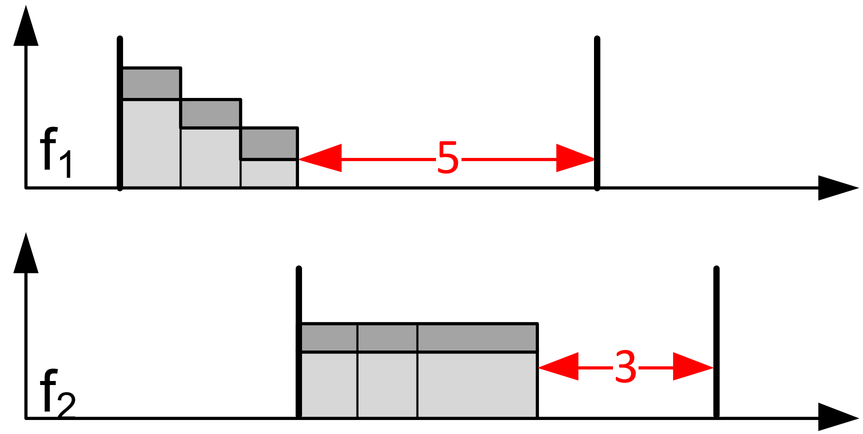
- Disaggregation ensure the balance of amounts

N-to-M aggregation

- **Grouping:** partition flex-objects into groups based on grouping parameter values being within given thresholds
- **Bin-packing:** further partition each group to satisfy aggregate constraints (count, total min/max,…)
- **N-to-1 aggregation:** as before, applied on every group



Time Flexibility Tolerance

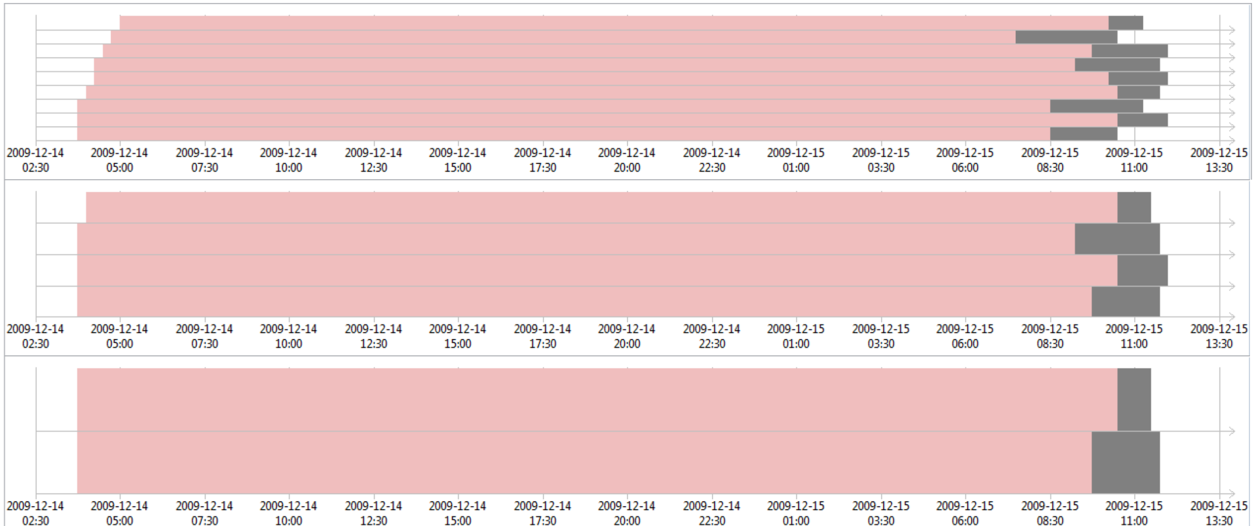


|5-3| ≤ TFT

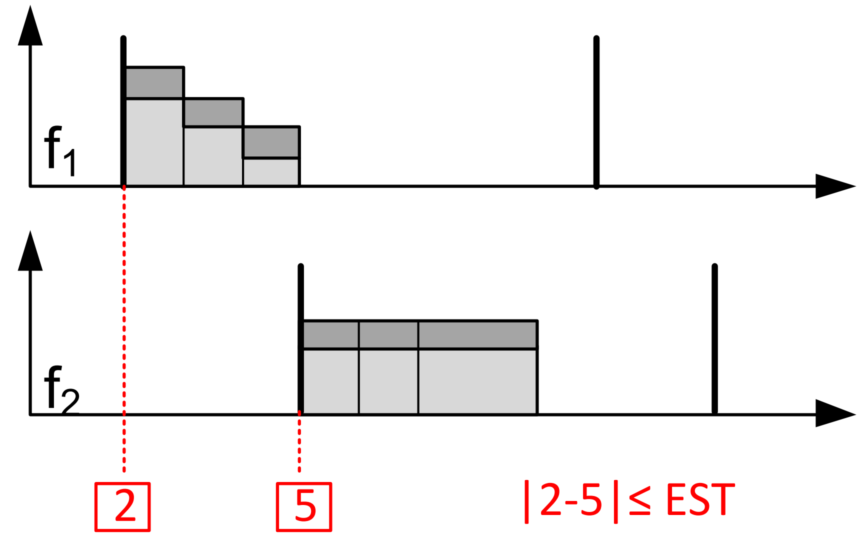
TFT=0

TFT=2

TFT=5



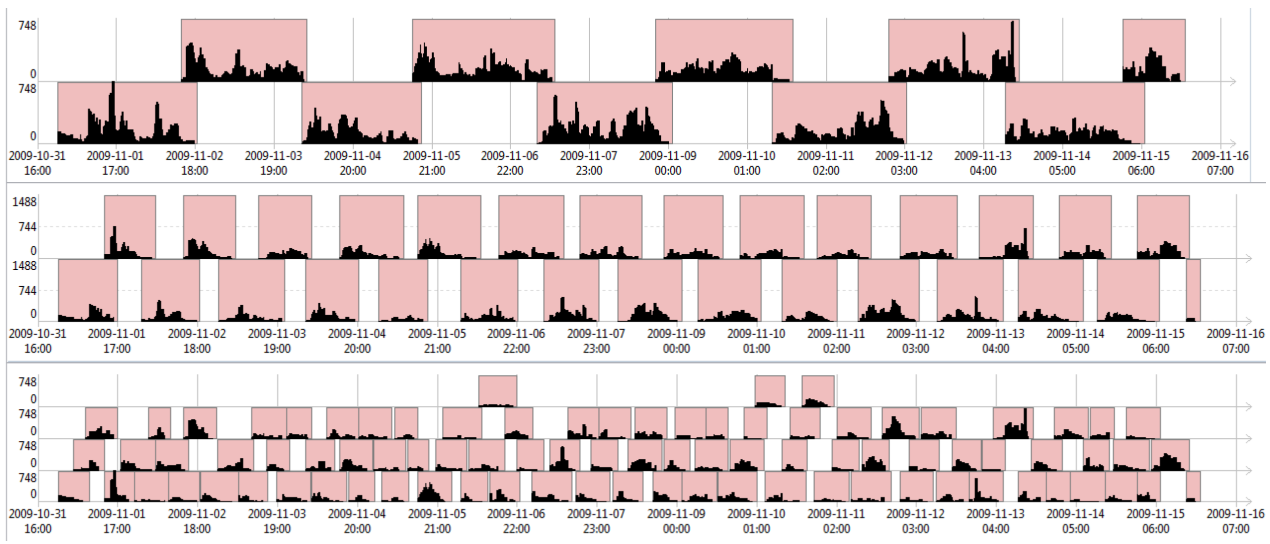
Earliest Start Tolerance



EST=150

EST=50

EST=15



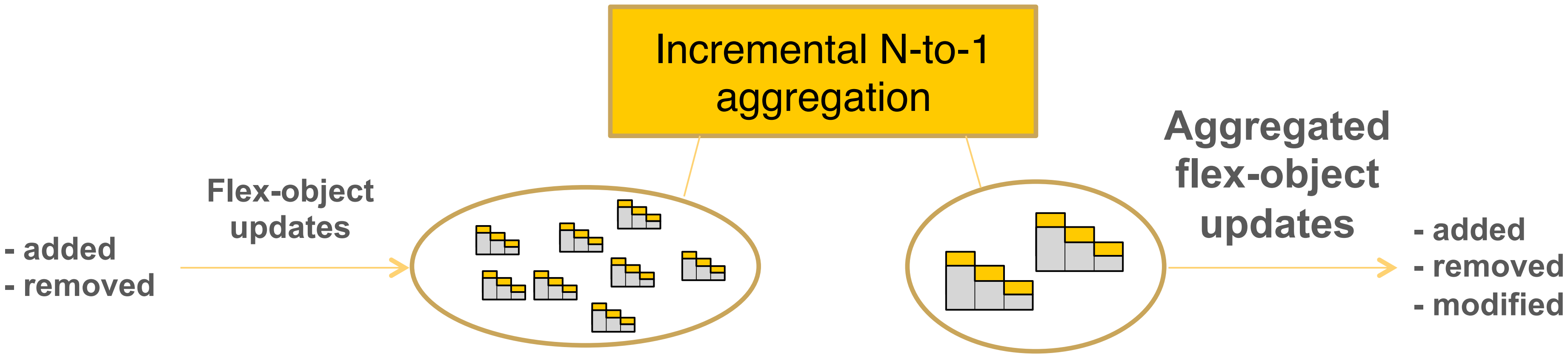
Parameter Settings

- The user will choose from a number of meaningful pre-defined parameter settings
  - Short/long profiles
  - Amount as early as possible
  - …

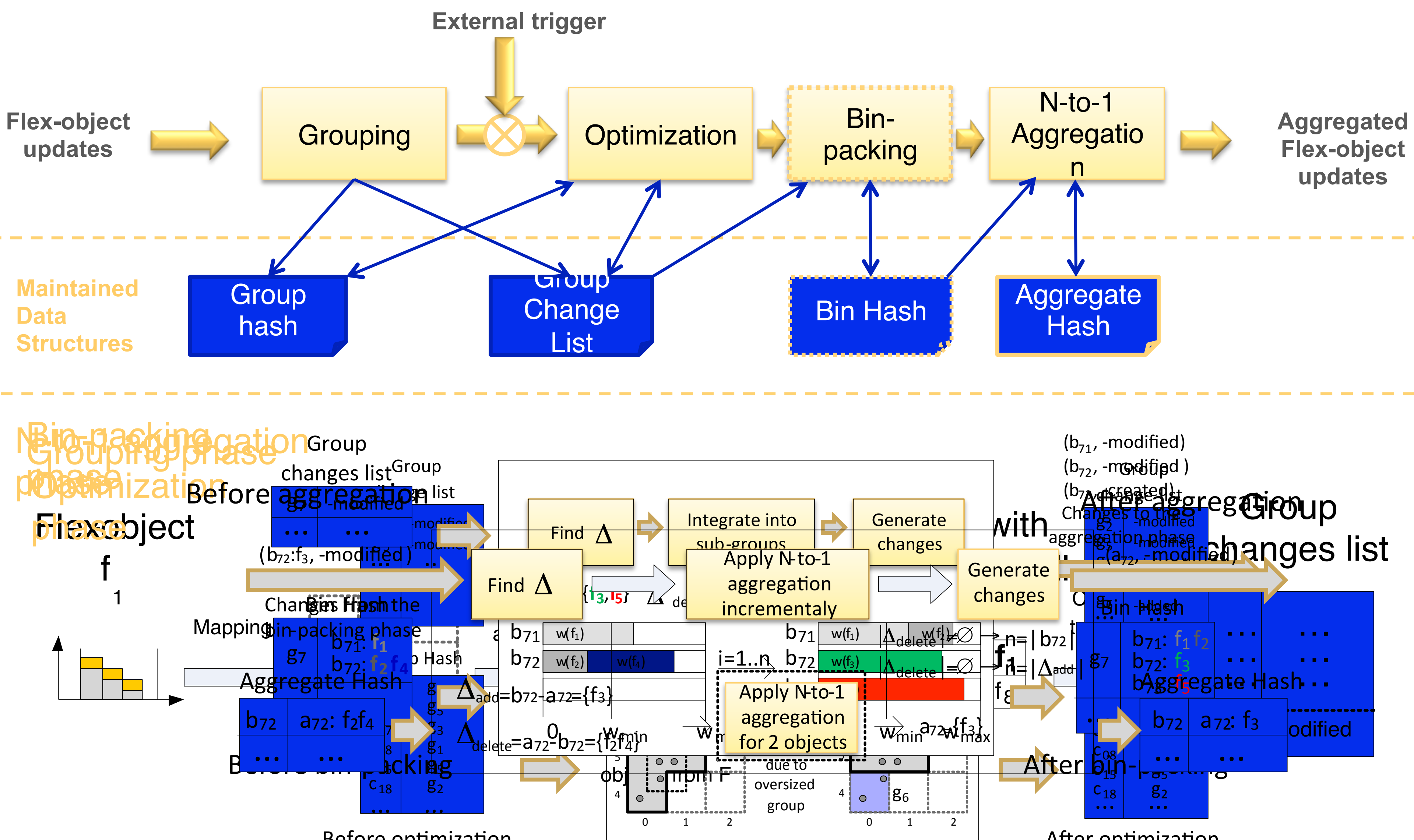
Incremental N-to-M aggregation

Main contribution

- Incremental grouping
- Incremental optimization
- Incremental bin-packing
- Incremental N-to-1 aggregation



Incremental N-to-M aggregation



Experimental Evaluation

Evaluation of the incremental N-to-M aggregation

- A synthetic flex-object dataset from the Mirabel project
- PC with Quad Core Intel R Xeon R E5320 CPU, 16GB RAM, OpenSUSE 11.4 (x86 64)

Experimental Evaluation

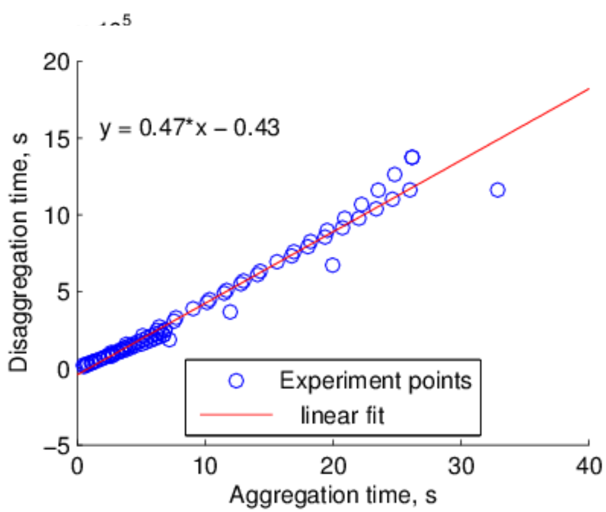
Scalability Experiment

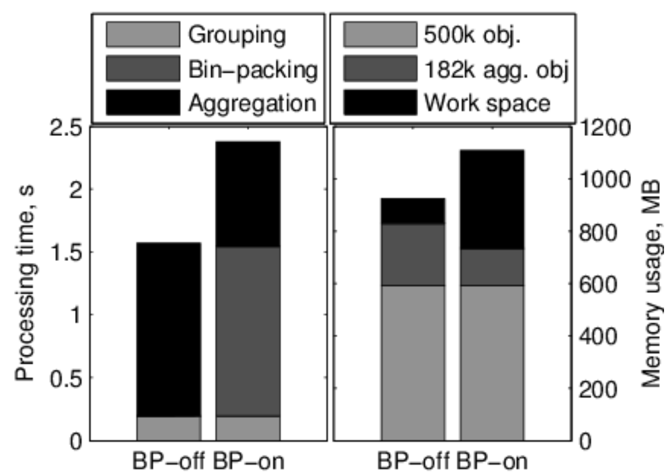
Variable Parameters

- Flex-object count: 50k … 1000k
- Grouping parameters
  - EST = 0, 250
  - TFT = 0, 6
- BP: enabled, disabled

Fixed Parameters

- BP ensures aggregated flex-objects with at least 2 hours of time flexibility





Experimental Evaluation

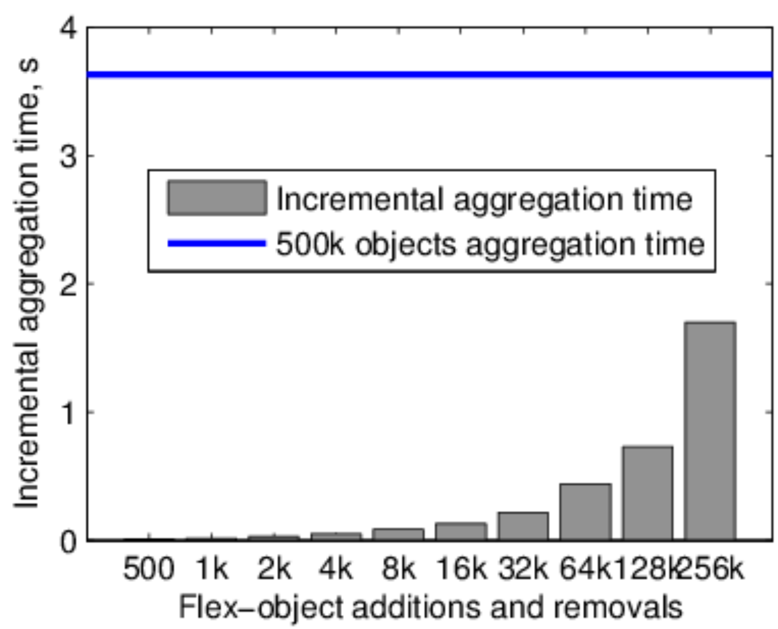
Incremental Behavior

Variable Parameters

- Number of inserts/deletes: 500..256k

Fixed Parameters

- Flex-object count: 500k
- BP: disabled
- Grouping parameters
  - EST = 0
  - TFT = 0



Experimental Evaluation

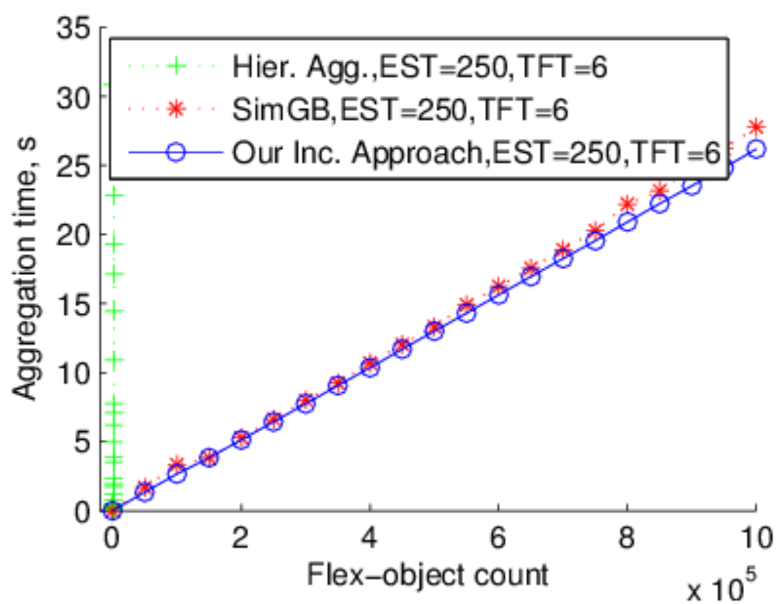
Comparison w. **partial** baselines (grouping only, non-incremental)

Variable Parameters

- Flex-object count: 50k … 1000k
- Partial baselines
  1. Hierarchical clustering
  2. Similarity Group By (Silva, et al.)

Fixed Parameters

- BP: disabled
- Grouping parameters
  - EST = 250
  - TFT = 6



Experimental Evaluation

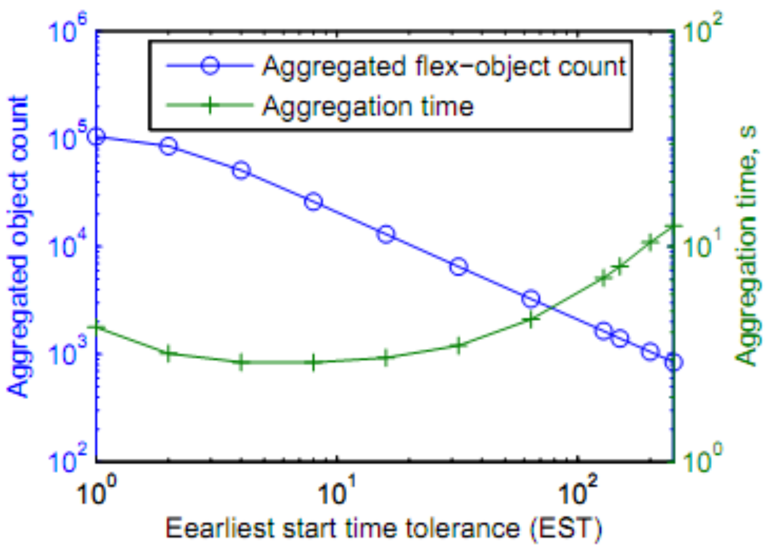
Grouping Parameter Effect

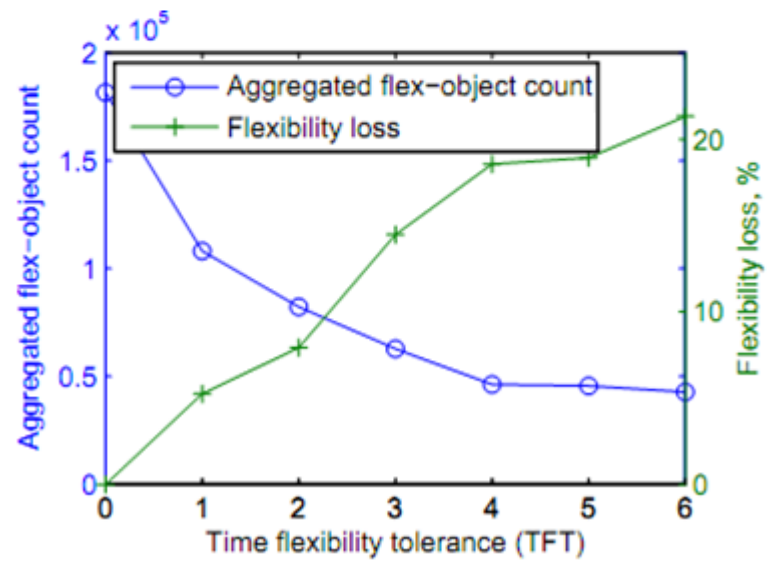
Variable Parameters

- Grouping Parameters
  - EST
  - TFT

Fixed Parameters

- Flex-object count: 500k
- BP: disabled





Experimental Evaluation

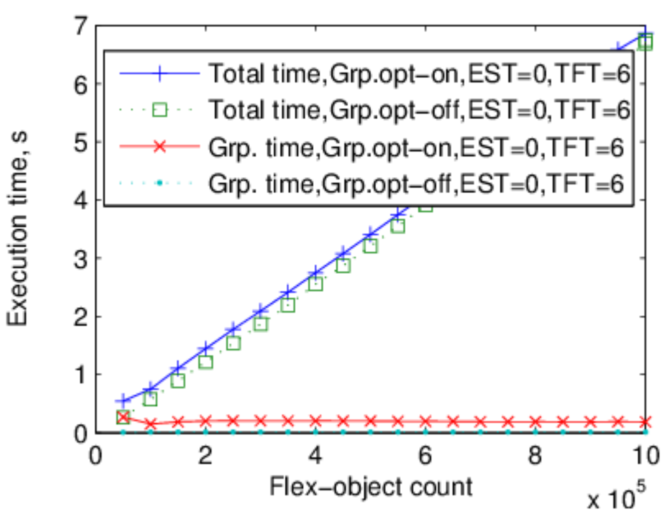
Group optimization phase effect

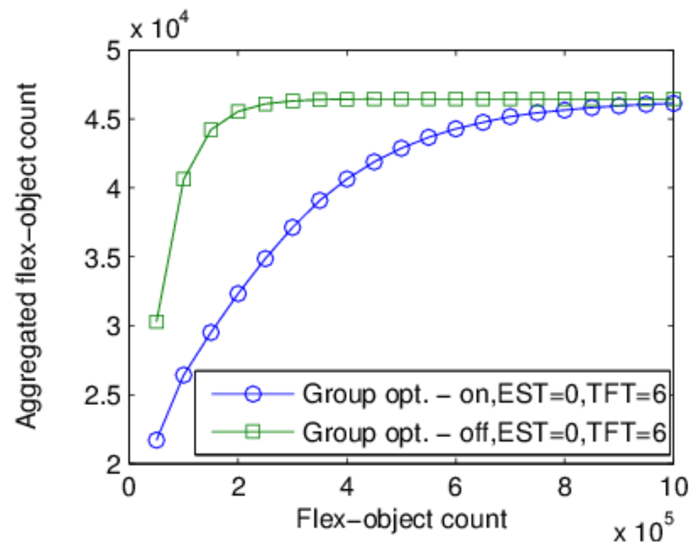
Variable Parameters

- Flex-object count: 50k … 1000k
- Group optimization phase: enabled, disabled

Fixed Parameters

- BP: disabled
- Grouping parameters
  - EST = 0
  - TFT = 6





Experimental Evaluation

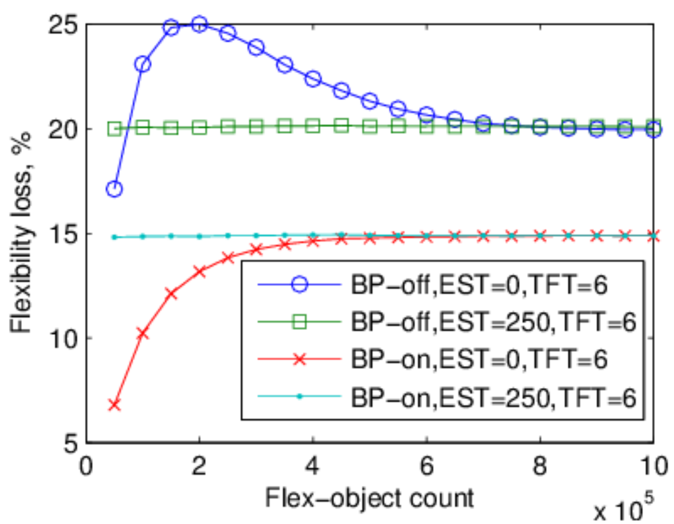
Bin-packing effect

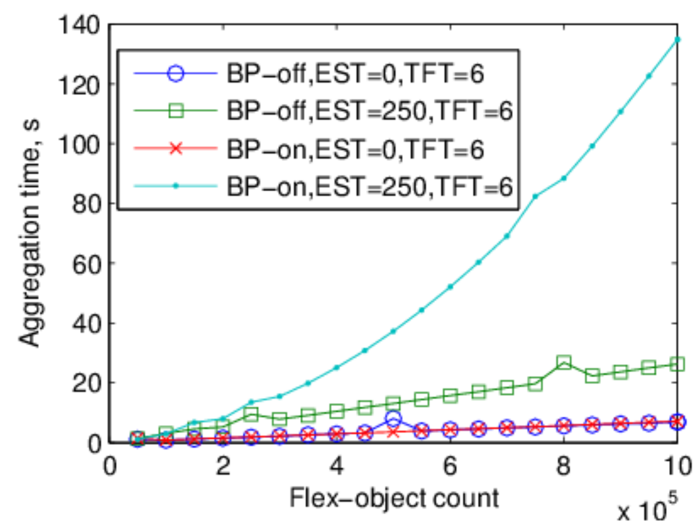
Variable Parameters

- Flex-object count: 50k … 1000k
- BP: enabled, disabled

Fixed Parameters

- BP ensures aggregated flex-object with at least 2 hours of time flexibility
- Grouping parameters
  - EST = 0, 250
  - TFT = 0, 6





Aggregation Conclusions

- Flex-objects allows planning of various processes, e.g., energy use
- A database handling flex-objects is needed
- Aggregation and disaggregation are two most important operations/queries
- Presented 3 aggregation techniques
- Experiments with the incremental N-to-M approach
  - Compression and performance of aggregation depends on grouping parameters
  - Aggregation and disaggregation can be done in linear time (BP-off)
  - When flex-object change marginally, incremental aggregation allows saving lots of aggregation time
  - Optimization step is effective
  - Grouping step is as fast as efficient non-incremental baselines

Aggregation Future Work

- Design the components of the flex-object database
  - Flex-object storage
  - Visualization
  - Techniques to process queries
  - Techniquse to optimize queries

- Support other types of flexibility

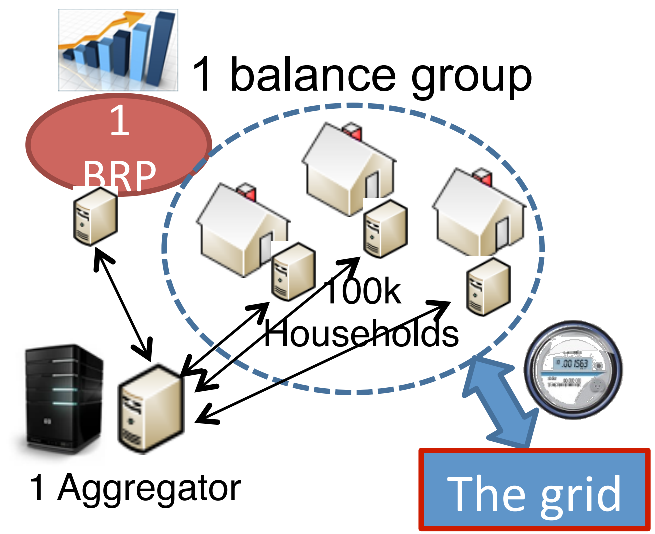
> Flex-Offer Aggregation Experiment

Situation today:

- BRP buys energy 24 hours in advance
- BRP is responsible for imbalances
- Imbalances are penalized

Our additions/scenario:

- Smart-grid CPS is introduced
- 1 household defines 1 flex-offer
- Flex-offers used for **consumption corrections**
- Flex-offers are available 1 hour before delivery
  - 10 minutes for *scheduling*
  - 50 minutes for *aggregation+disaggregation*



Experiment

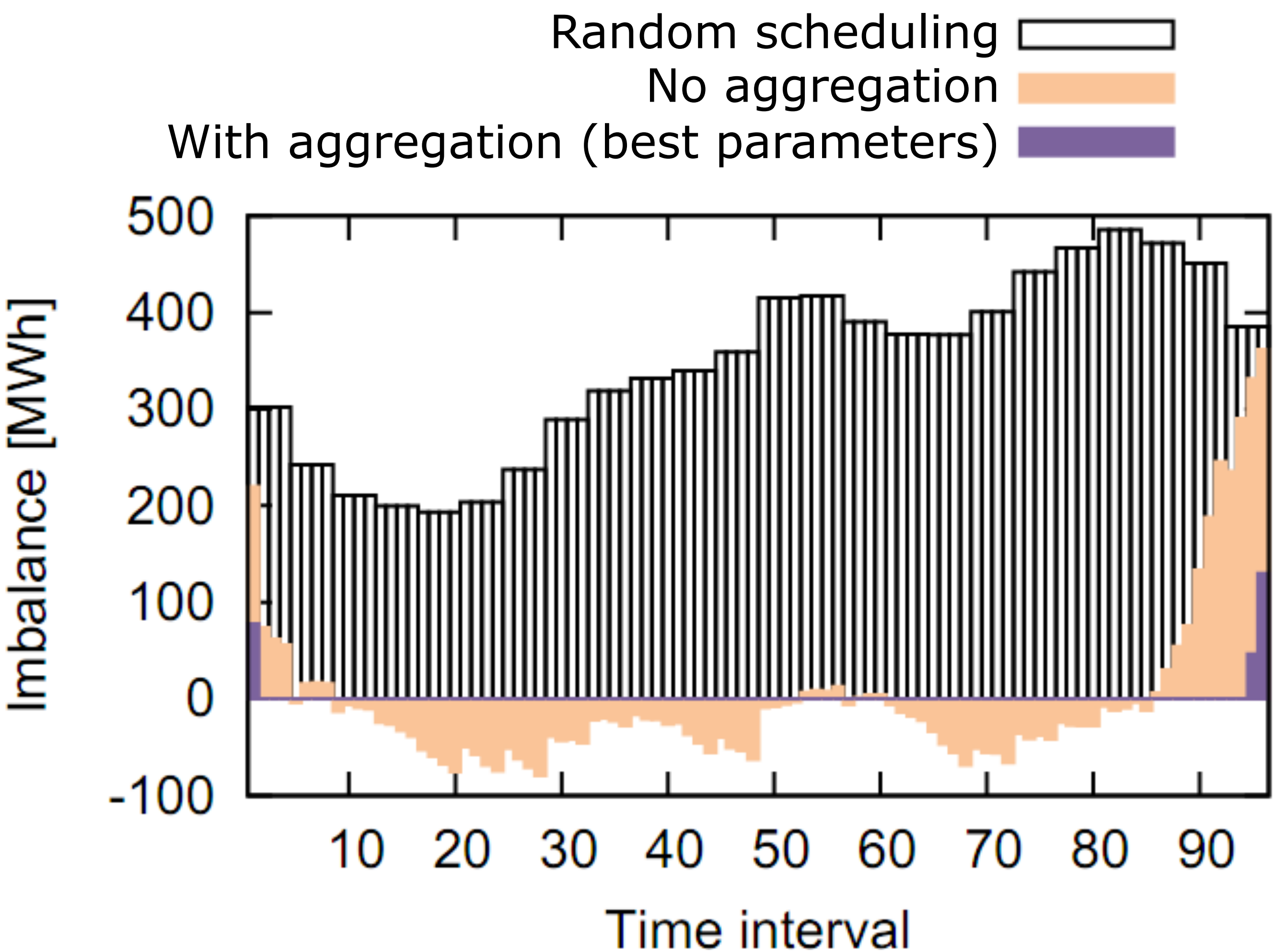
- Generate 100k flex-offers based on real data
- Use real energy prices from Slovenia
- Day ahead schedule has "correct amount", but amount is "incorrectly distributed"

BRP minimizes the cost function:

$c =$ Cost of remaining imbalances + Cost of flex-offer assignments + Cost of energy to be bought or sold on the market

> Flex-Offer Aggregation Experiment

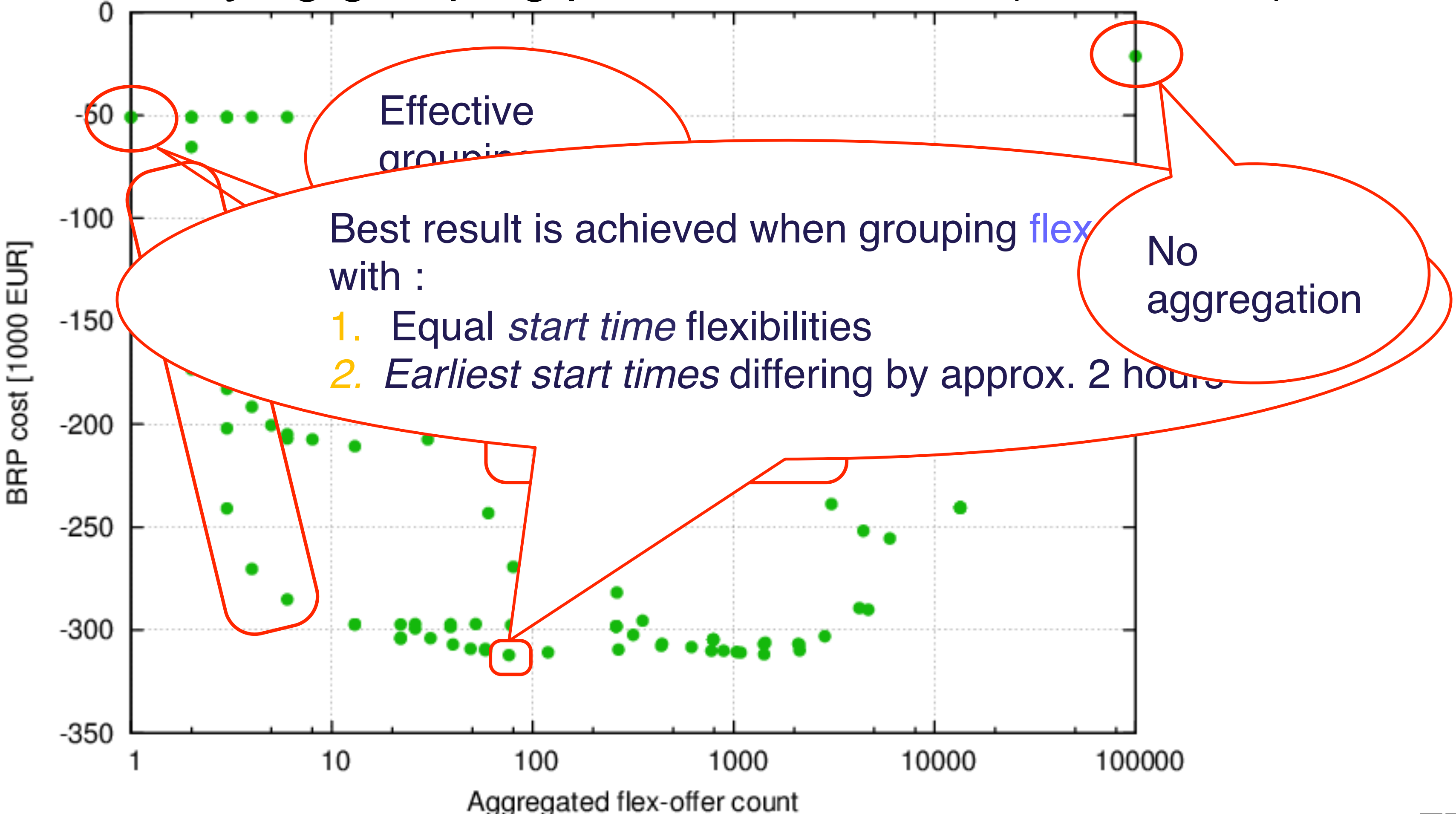
➢ FLEX-OFFER-BASED BALANCING (MAX 10 MIN FOR SCH) IMBALANCES IN THE BALANCE GROUP (MWH)



> Flex-Offer Aggregation Experiment

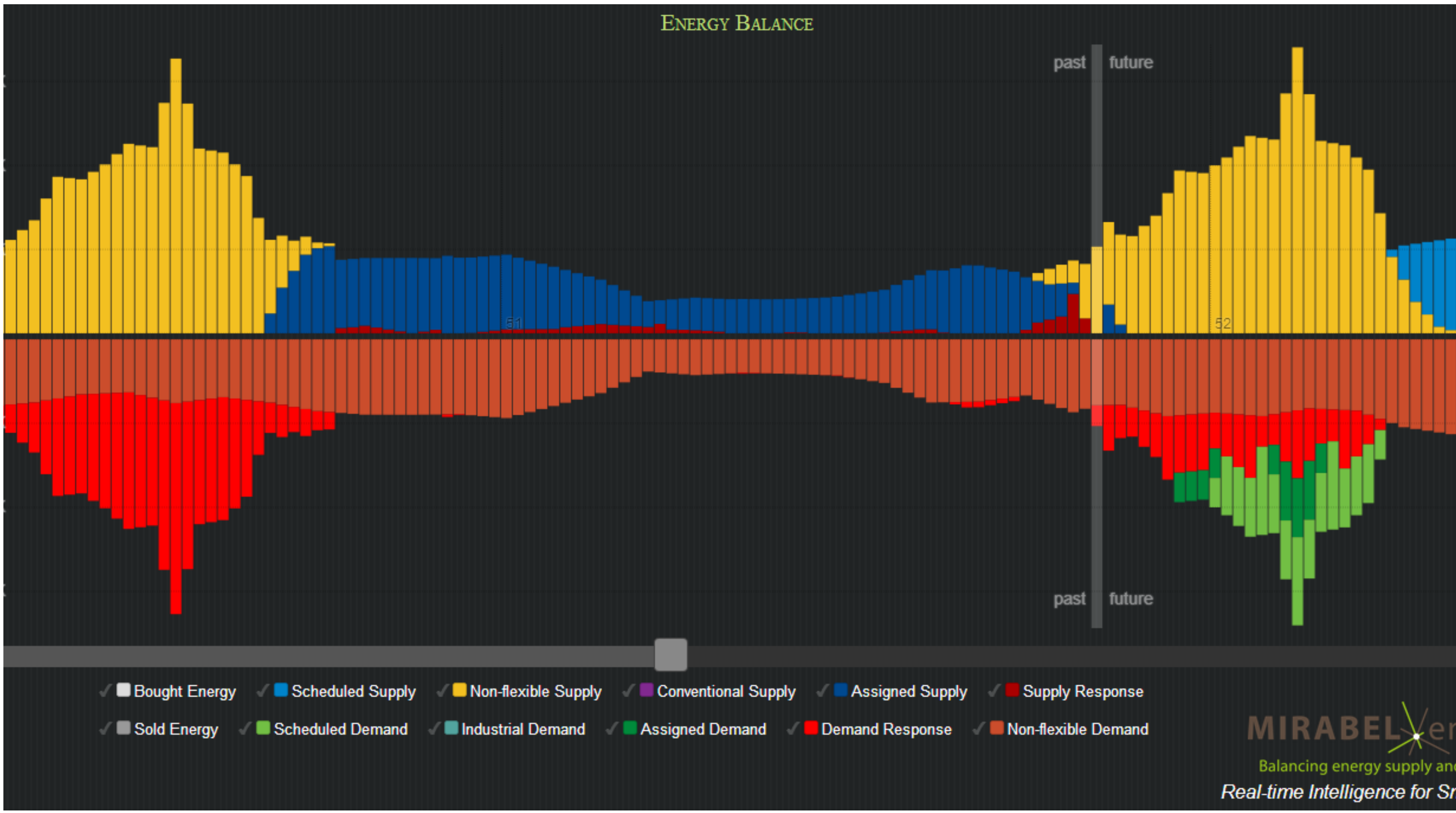
➢ THE EFFECTS OF FLEX-OFFER-BASED BALANCING

BRP costs **with** and **without** aggregation (reduce[A]) while varying grouping parameter values (1000 EUR)

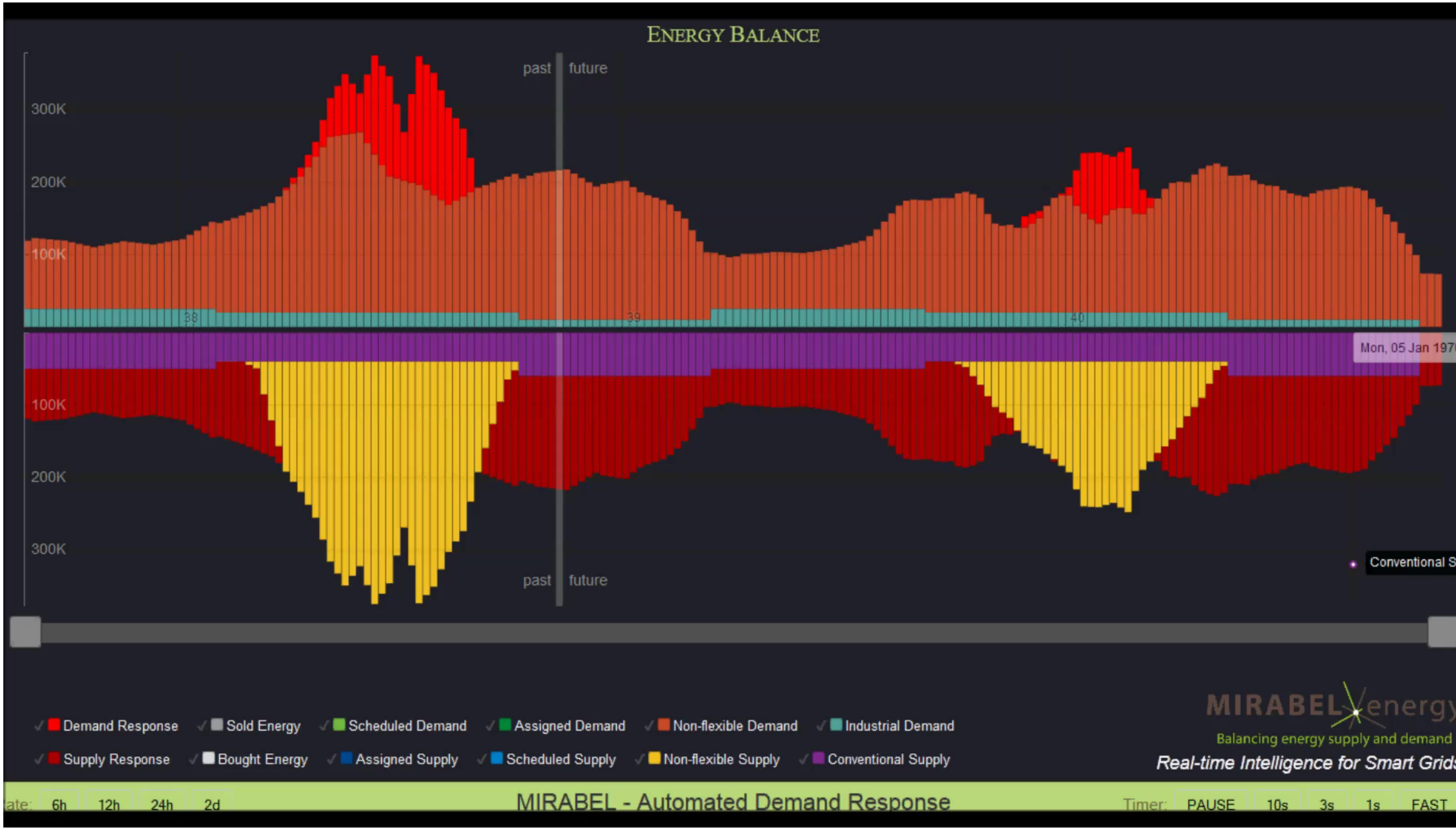


MIRABEL Prototype

Balancing electricity supply and demand in near real-time



MIRABEL In action



MIRABEL Experimental Results

- 7-13% BRP **cost** reduction
- 13-50% **peak-load** reduction
- Increase of **base-load**
- Improving RES **integration** significantly
  - 70% of the negative impact of fluctuating renewables can be **neutralized** if 15% of the energy consumption is flexible and intelligently controlled by the BRP.
- Households can reduce energy **bills** by 10-20%.
- With energy storage: up to 50% **CO^2** reduction
- Aggregation+scheduling better+faster than just scheduling
- Even better with less conservative flexibility assumptions

Ongoing Project: Totalflex

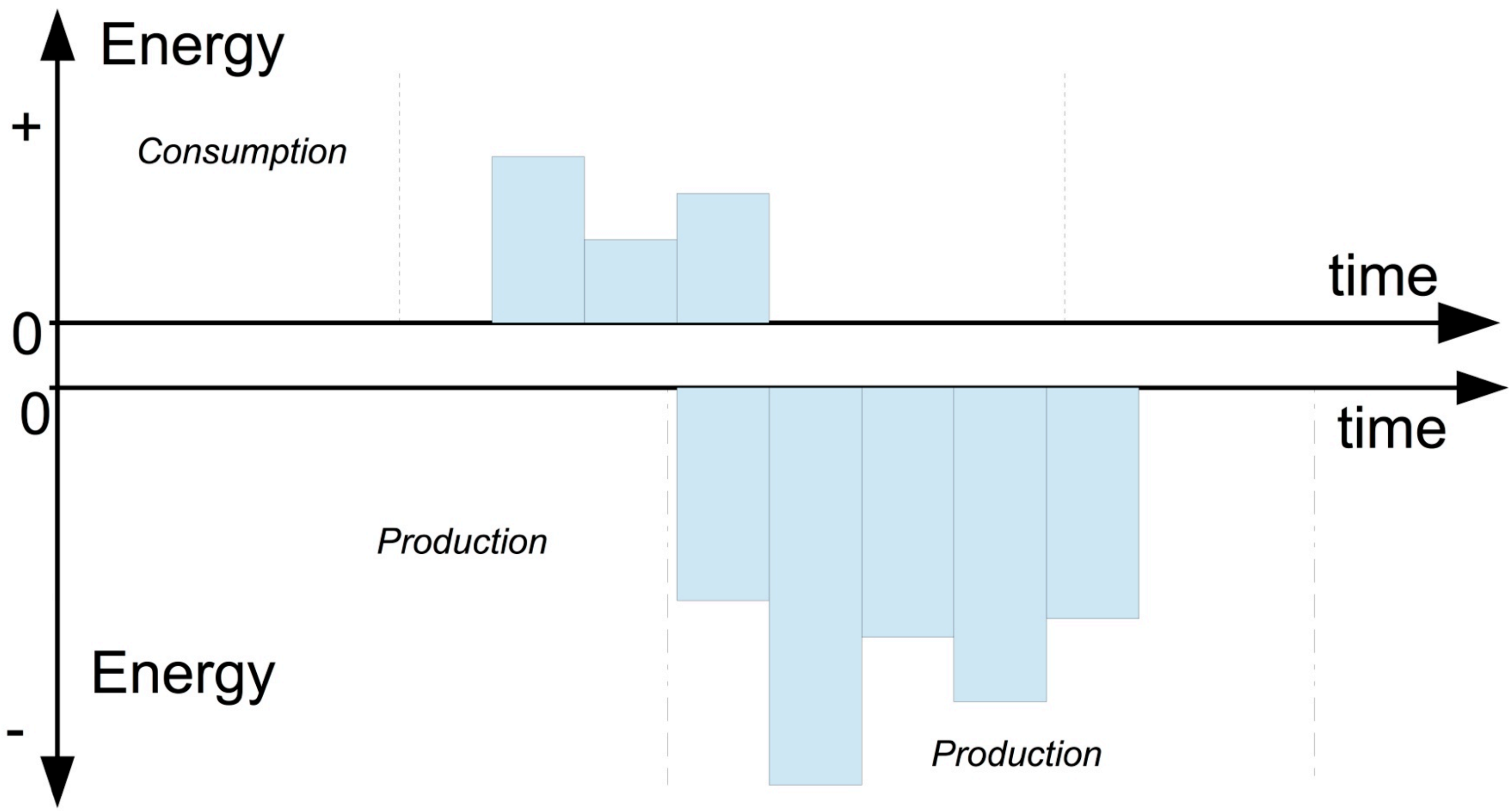
- *"The vision of TotalFlex is to develop a cost-effective, market-based system that utilizes total flexibility in energy demand and production, taking balance and grid constraints into account"*
- www.totalflex.dk
- Extending MIRABEL **downwards** into the home…
  - Home automation integration: device level measurements/control
  - Prediction of consumption/flexibility at *device* level: auto-gen Fos
- …and **outwards** to capture more aspects
  - More advanced FO aggregation and analysis
  - Modeling heatpumps, etc., as FOs
  - Balancing demand and supply in more aspects
  - Help DSO distribution grid management, e.g., avoid congestions

Totalflex Balance Aggregation

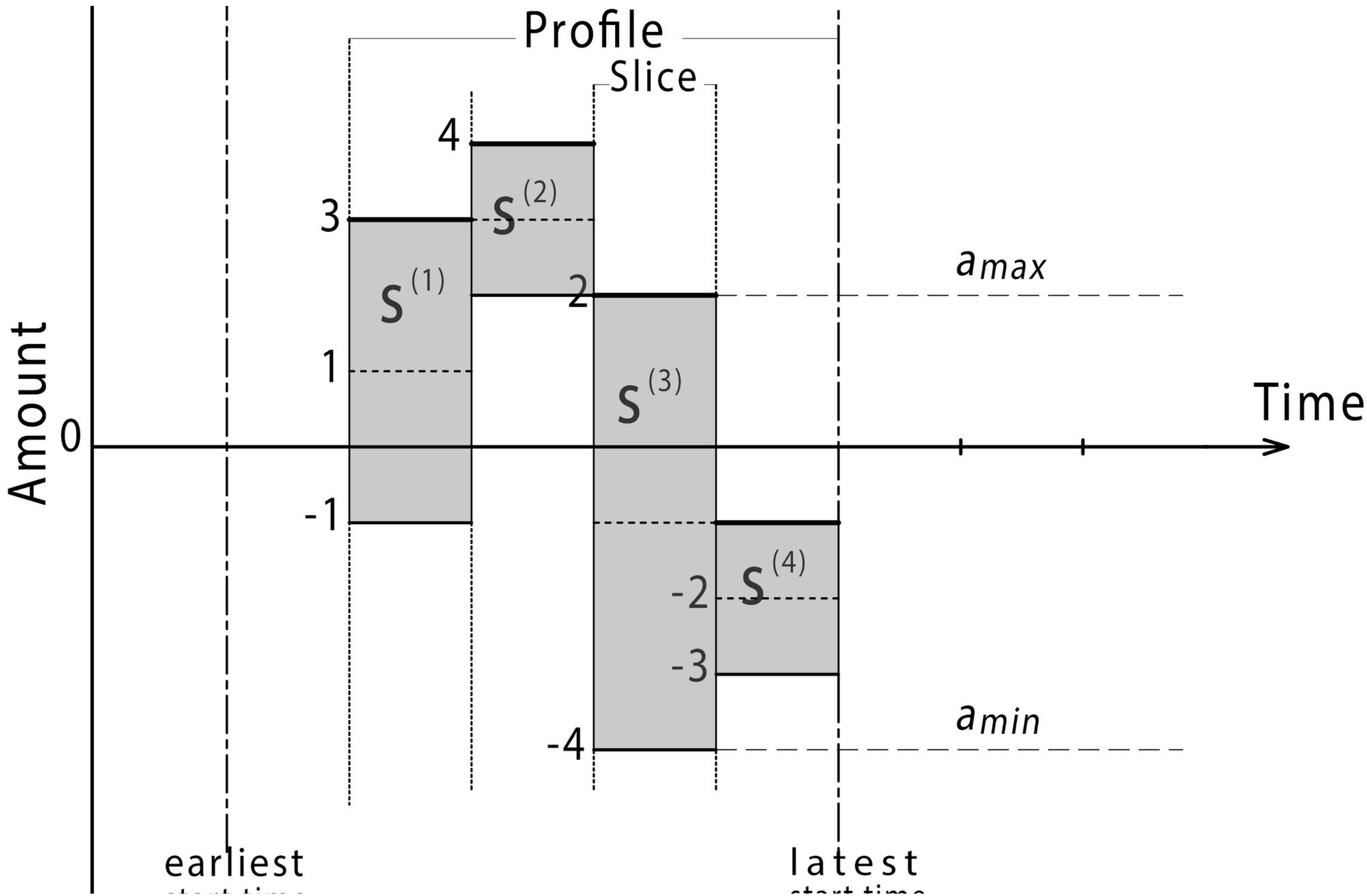
- Initial flex-offer aggregation considers:
  - Only the **market** dimension
  - Not the **physical grid** dimension
- A large flex-offer can **violate** local capacity constraints
  - Perhaps in combination with other flex-offers
  - Example: charging several EVs in a single street – not enough spare capacity
    - Can cause **black-out (too little power, frequency drops)**
  - Reverse example: getting excess solar power out of a rural area
    - Can cause **white-out (too much power, frequency rises)**
- Observation
  - Supply and demand can (partly) **cancel out** each other locally
- Idea:
  - Aggregate flex-offers together to achieve (local) balance

Flex-offers

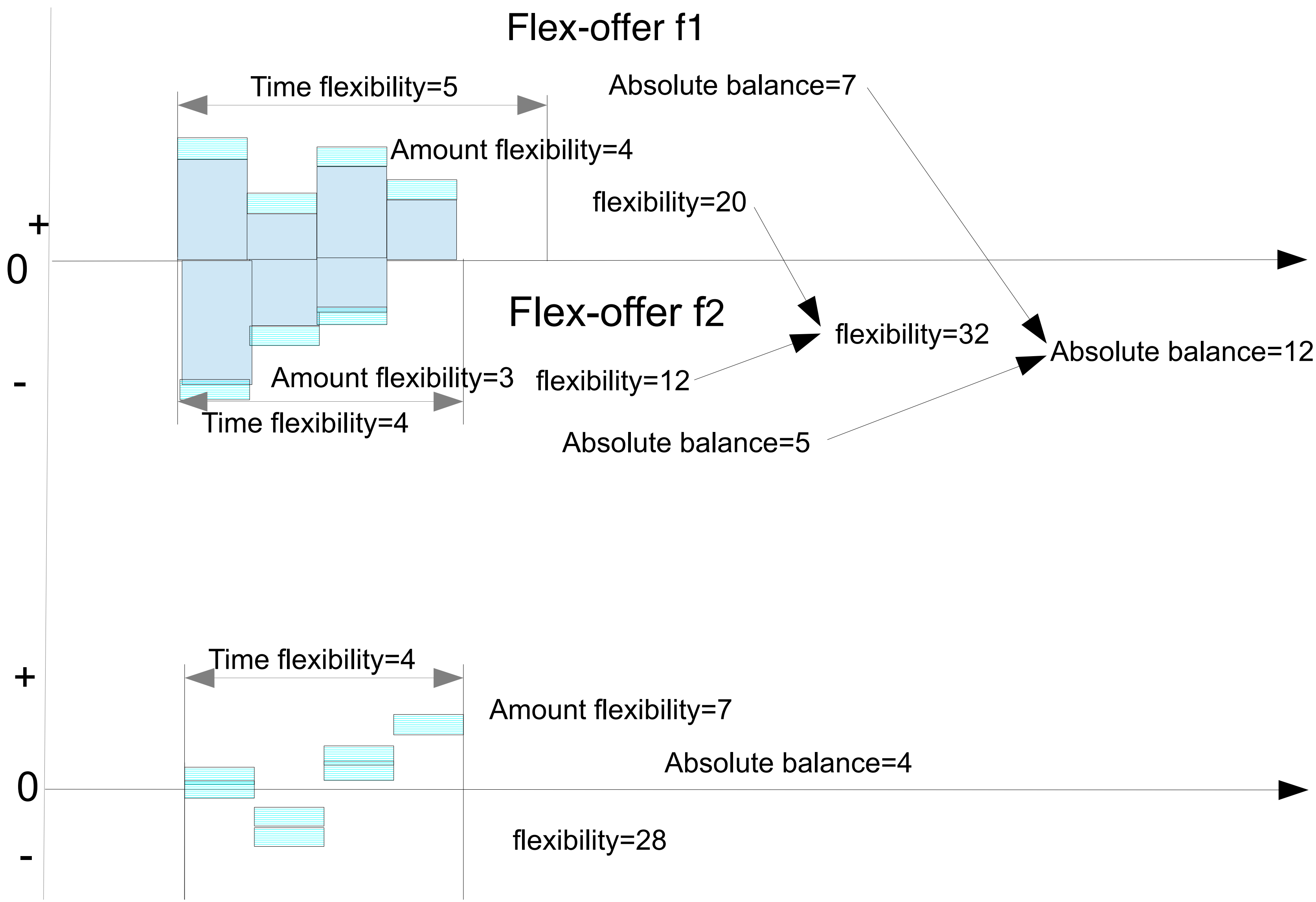
- Positive – Negative



Mixed flex-offer



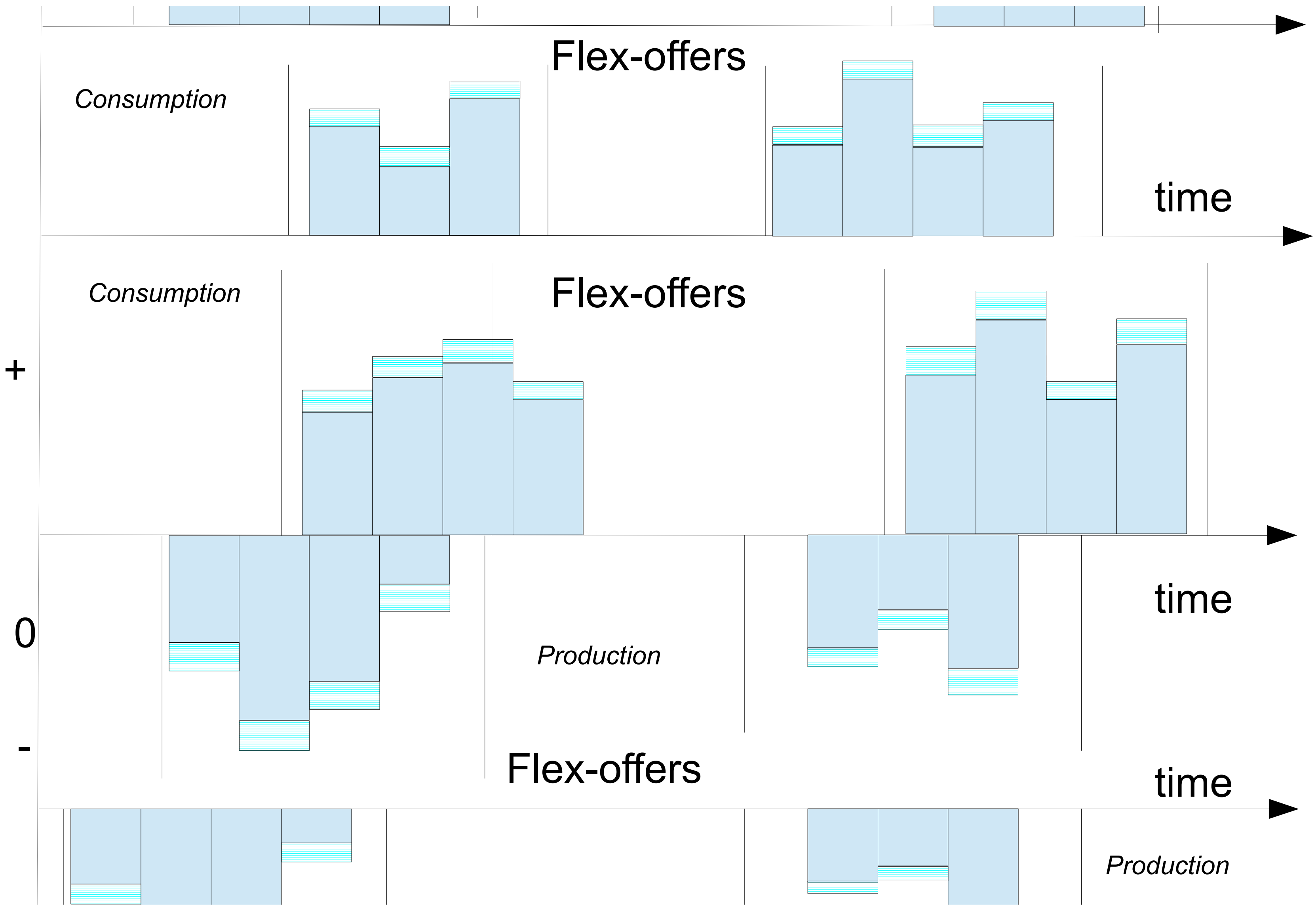
Flexibility loss example



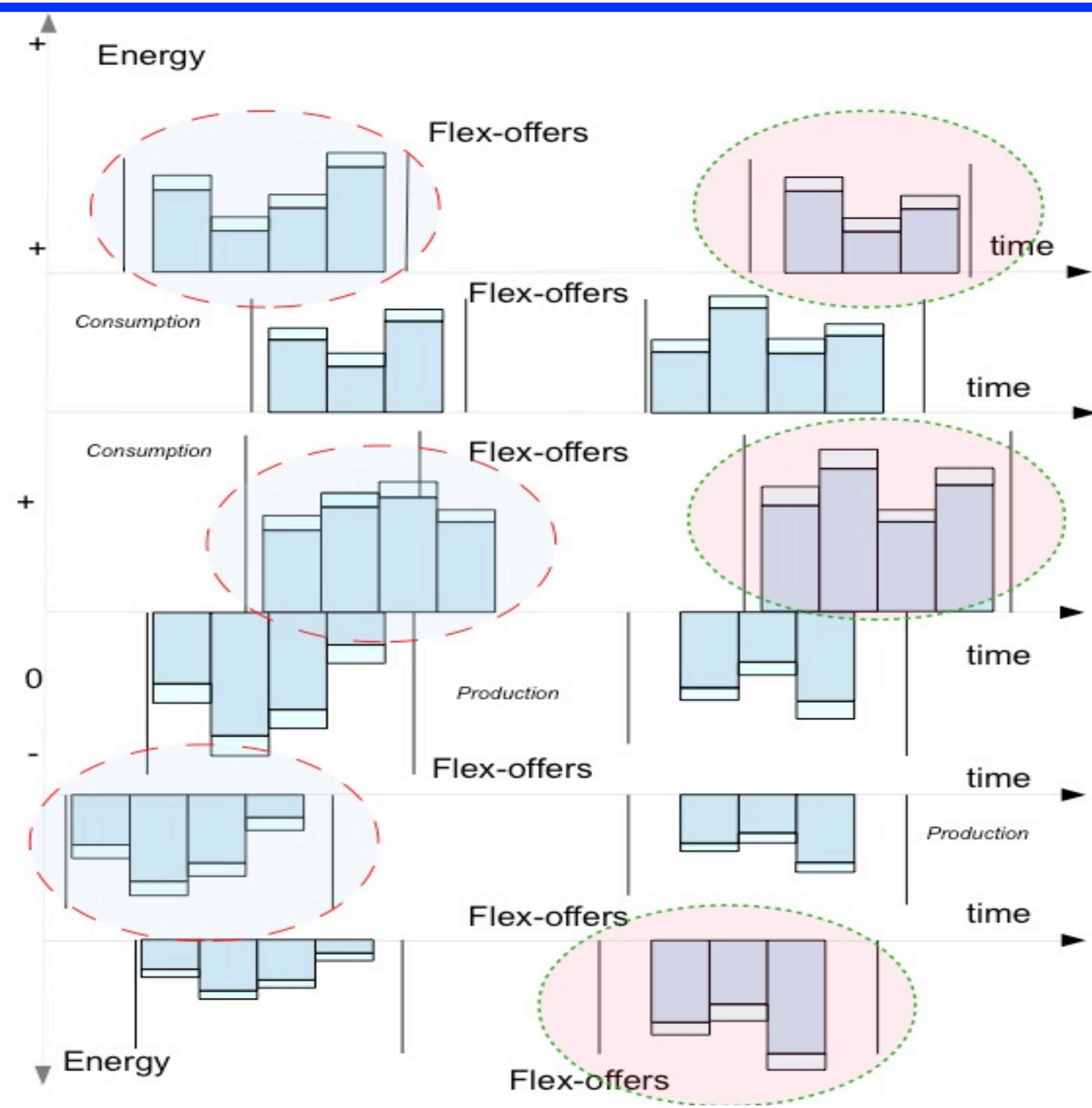
Why do we aggregate flex-offers?

- Trade on the market macro flex offer
- Reduce the planning complexity
- Stable electricity grid
- Handle imbalances
- Provide anonymity

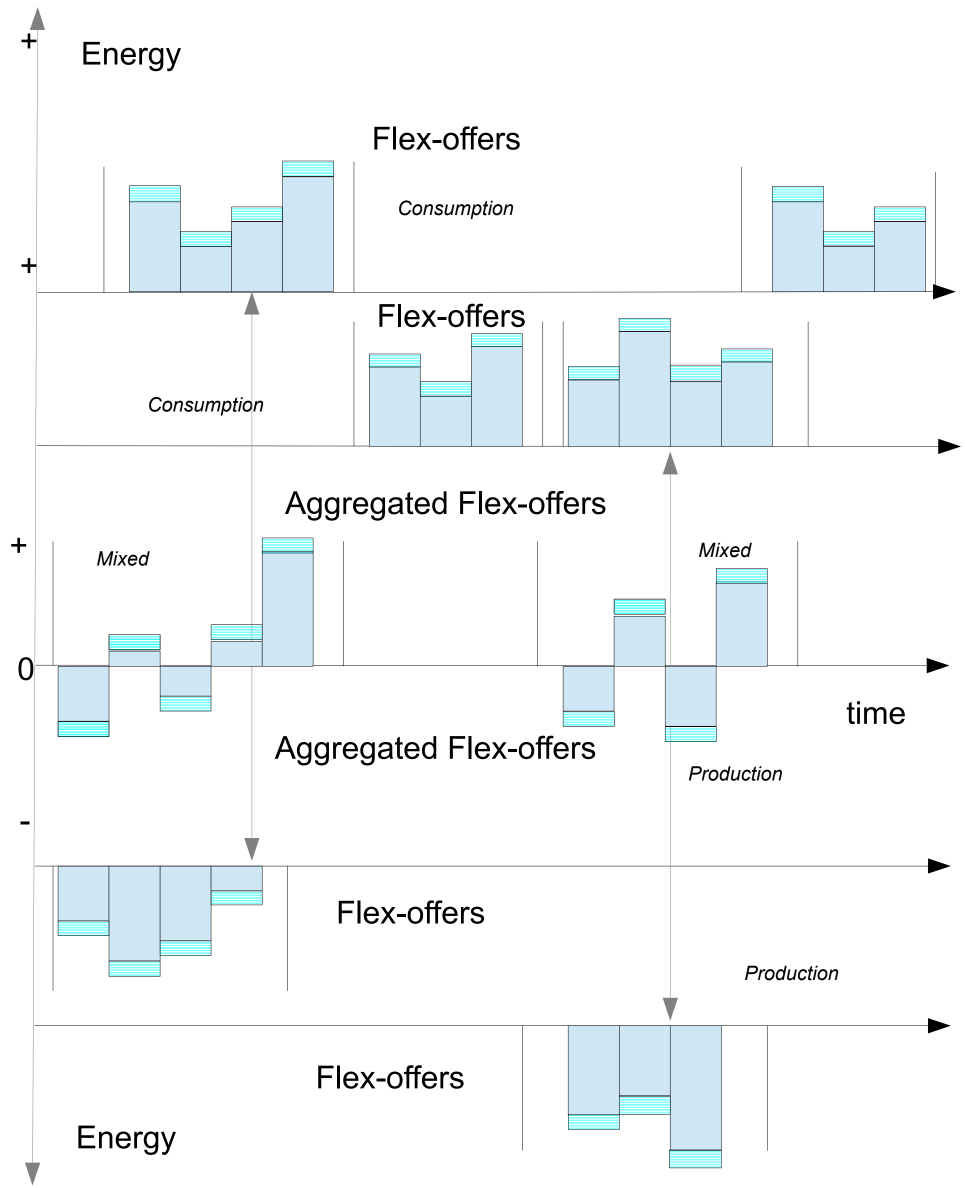
Balance aggregation - Input



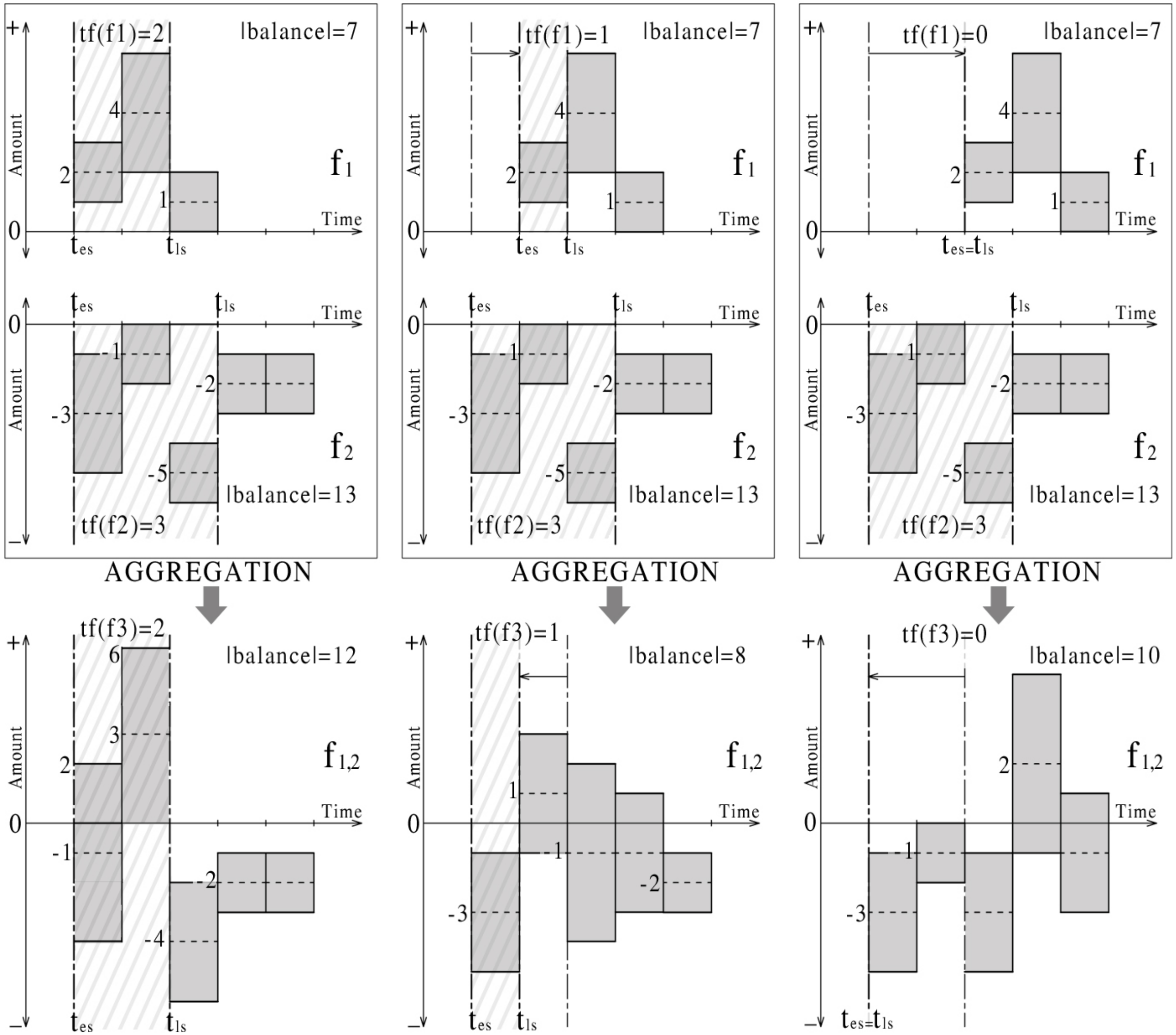
Balance aggregation - Grouping



Balance aggregation - Aggregate



Balance aggregation example



Simple greedy approach

- Pick largest (negative) flex-offer of the group, f_min
- balance = GetBalance(f_min)
- Iterate within the group
  - Pick the f with balance equal/closest to –balance
  - Aggregate with the currently aggregated
  - Update balance
  - Stop when balance is no longer reduced
- Start again in the same way

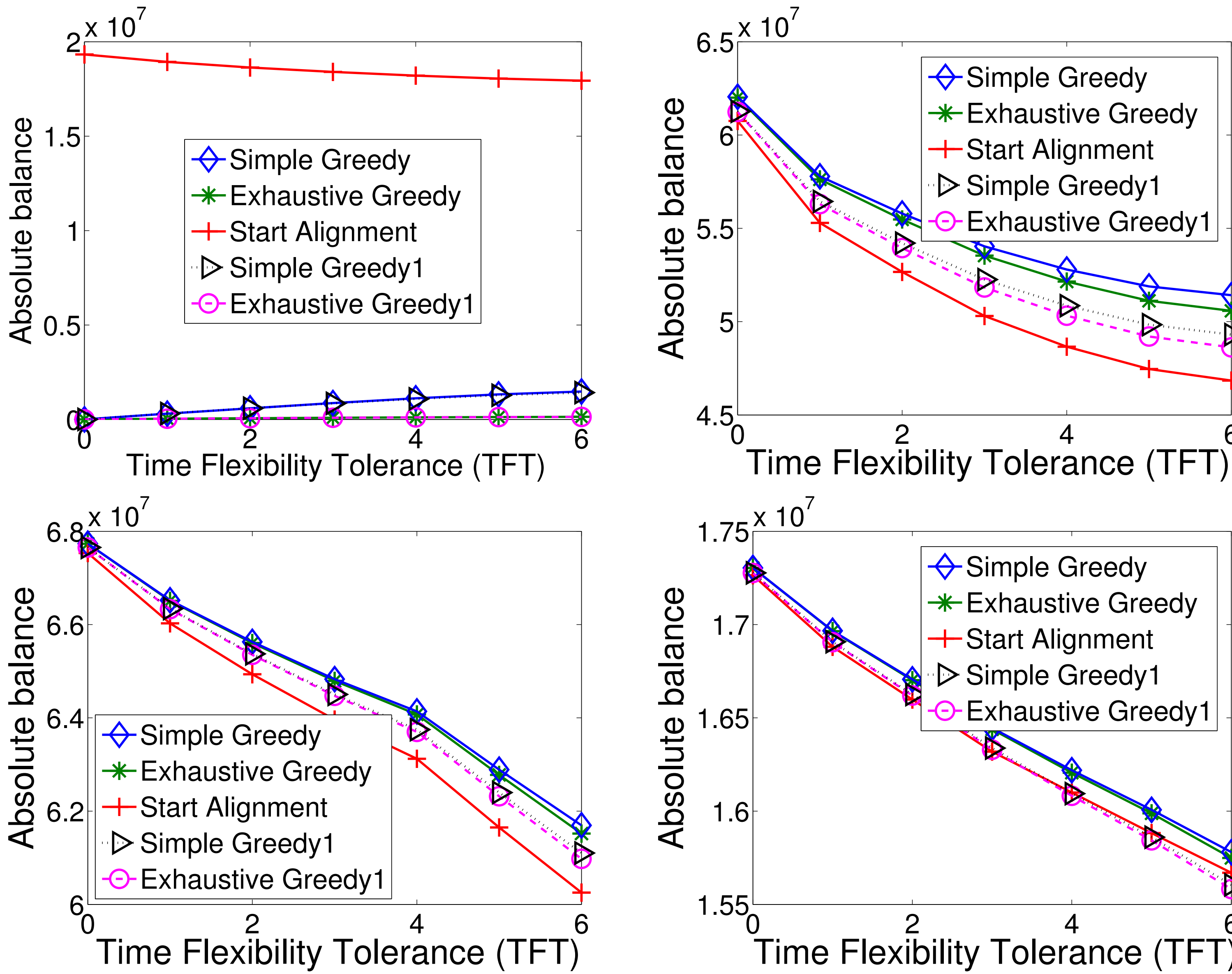
Exhaustive greedy approach

- Pick largest (negative) flex-offer of the group, f_min
- balance = GetBalance(f_min)
- Iterate within the group
  - Try all combinations
  - Pick the ONE that reduces the absolute balance the most
  - Update balance
  - Stop when balance is no longer reduced
- Start again in the same way

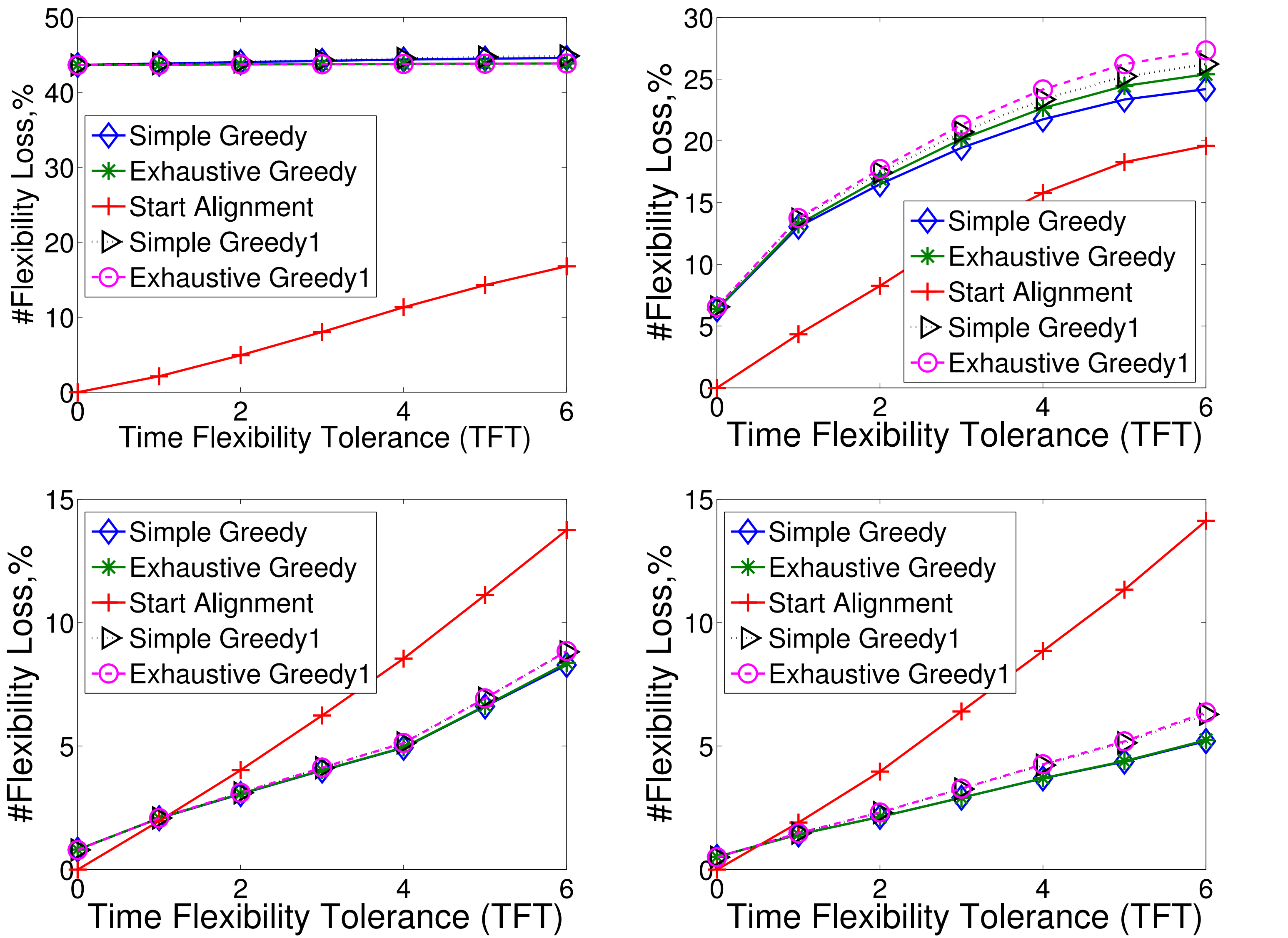
4 experimental setups

- 1st setup
  - Profiles from 2.5 to 7.5 hours long (production and consumption)
  - Time flexibility from 1 to 3 hours (production and consumption)
- 2nd setup
  - profiles are from 2.5 to 7.5 hours long (consumption)
  - Double length profiles for consumption
  - Half number for production flex offers
  - Same time flexibility between production and consumption
- 3rd setup
  - profiles are from 2.5 to 7.5 hours long (consumption)
  - Double length for production
  - Half number for production flex offers
  - Less time flexibility for production
- 4th setup
  - profiles are from 2.5 to 7.5 hours long (consumption)
  - More than double length profiles for production
  - Half number of production flex offers
  - Less time flexibility for production
  - Less energy flexibility for production

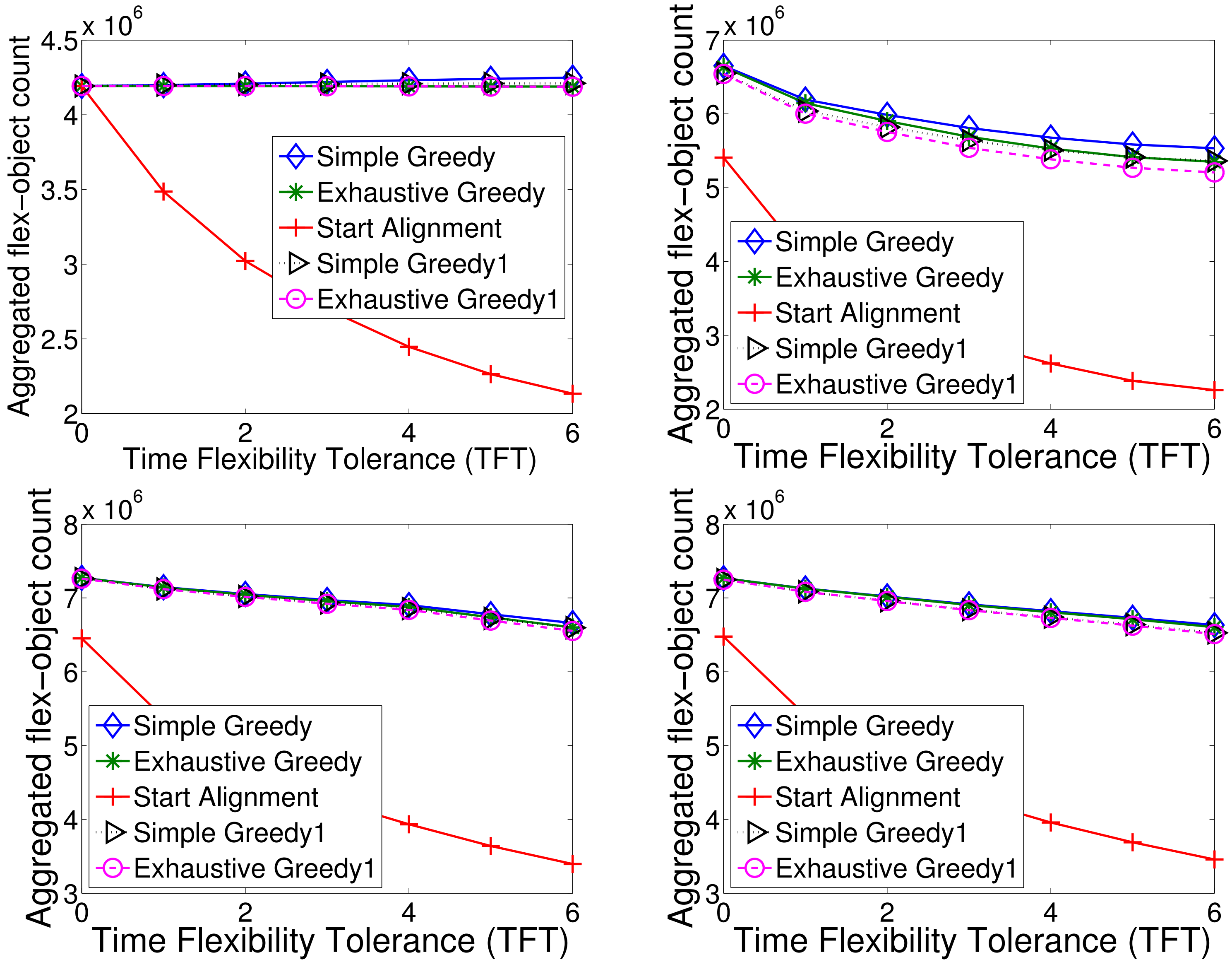
Absolute balance results



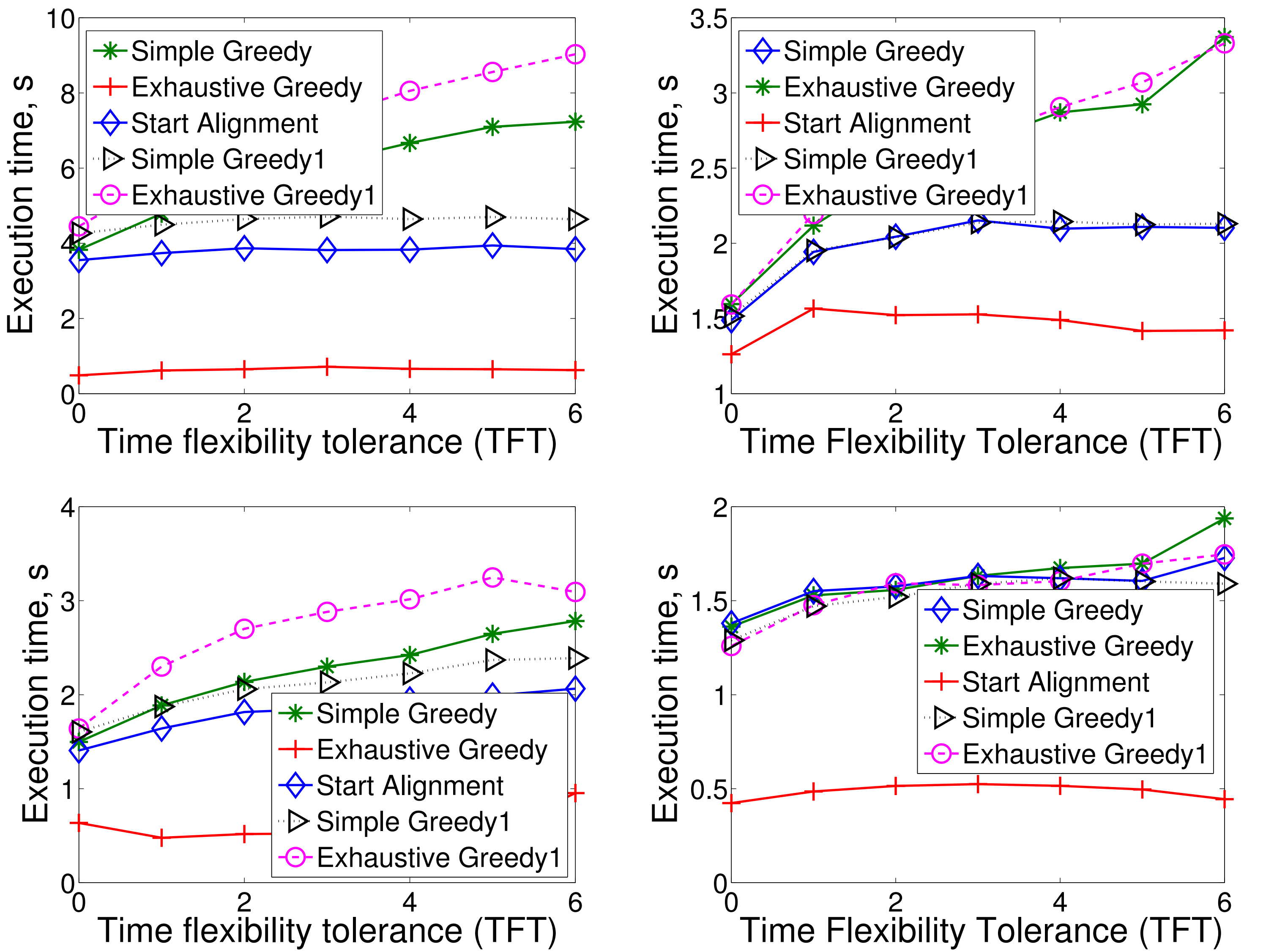
Flexibility loss results



Aggregated flex-offers counts



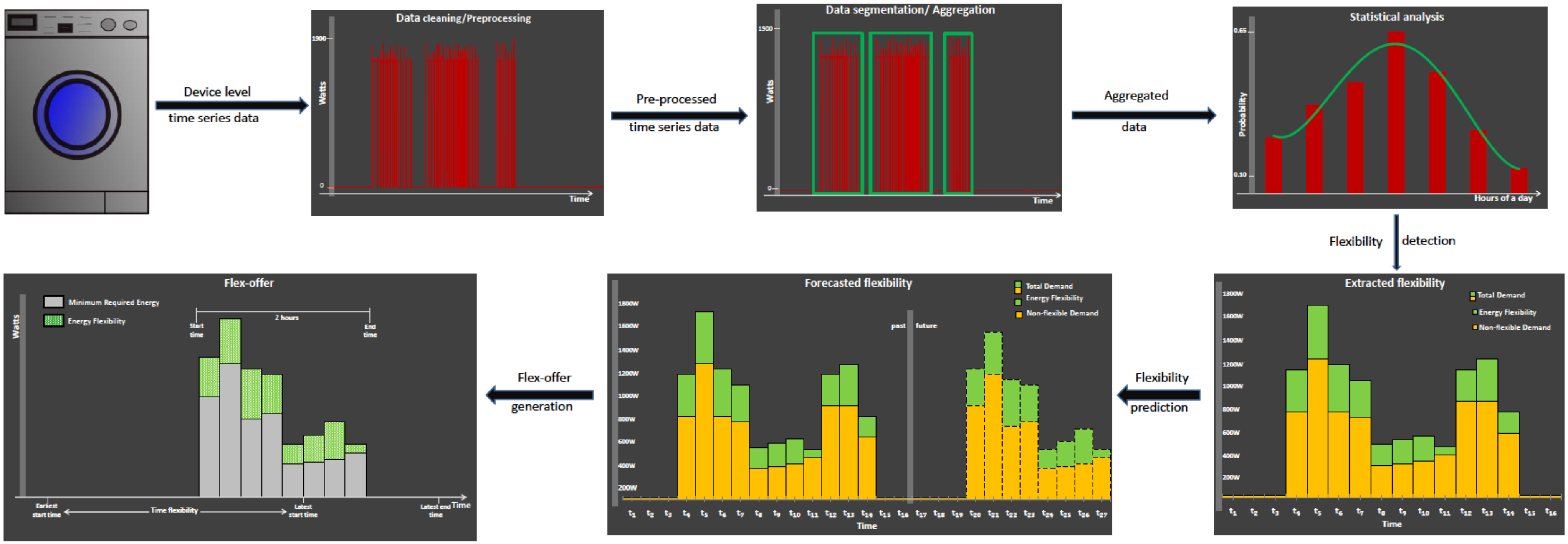
Execution time results



Balance Aggregation Summary

- Balance aggregation is feasible
  - We can achieve low balance (if possible)
- However, there is a tradeoff
  - Between balance, flexibility loss and processing time
- To get good balance
  - Sacrifice some time flexibility
  - Use more processing time
- The best technique depends on the scenario
  - For some scenarios, start aligment works well
  - For others, simple/exhaustive greedy works better

Totalflex Flexibility Forecasting

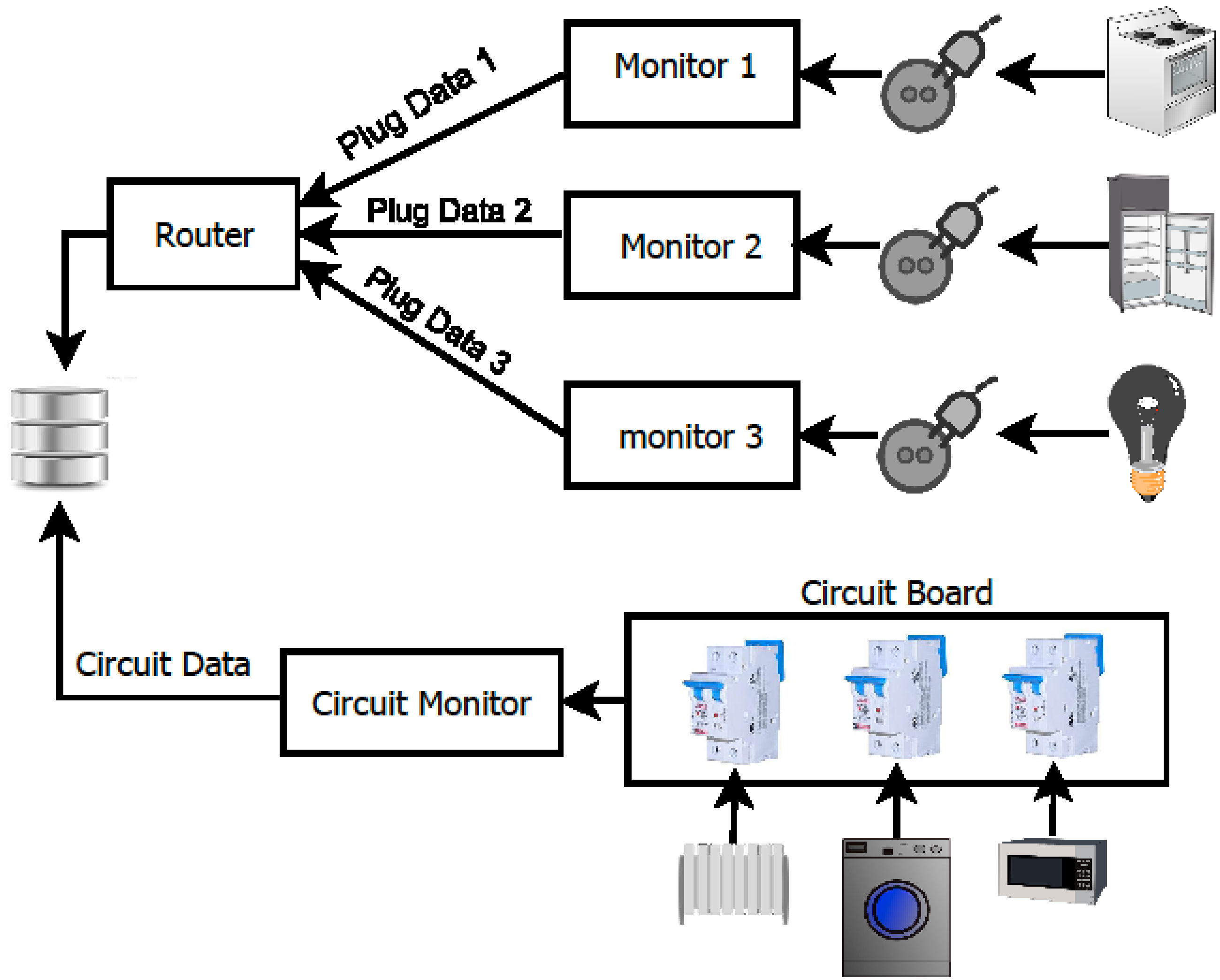


Flexibility Detection Study

- Initial study
  - Flexibility analysis and detection in device level data
  - Based on the North American REDD dataset
  - Totalflex device level data (smart sockets) being collected s

- Analysis on device level energy consumption.
- Device flexibility analysis.
- Users' device operation behaviors and patterns.
- A comprehensive device level analysis of energy consumption data.
  - Which will form the foundation for accurate flex-detection, flex-prediction, load-prediction.

REDD Data Collection



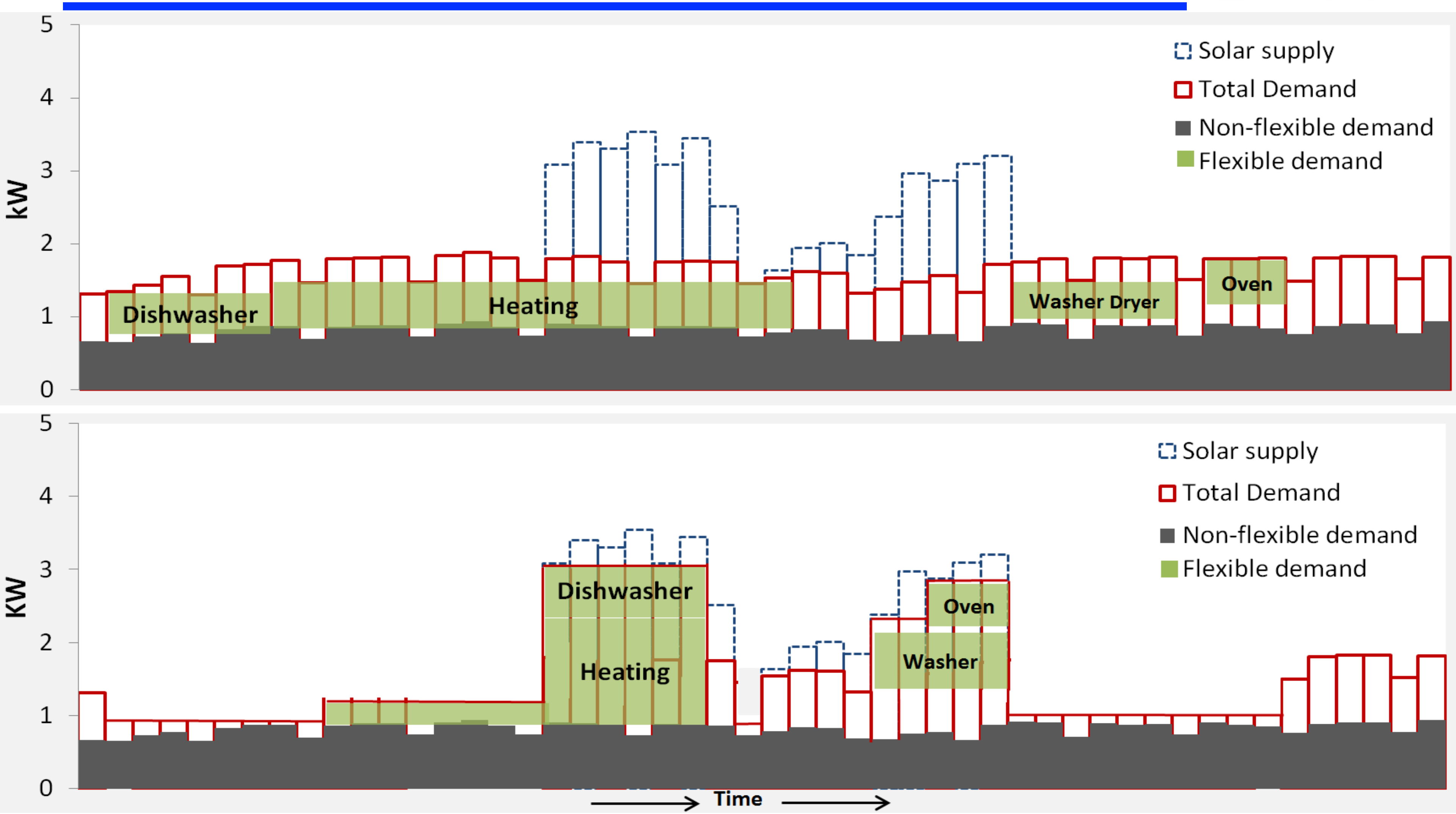
The REDD hardware architecture for data collection (adapted from REDD [4]).

Background

- Flexibility is in two dimensions:
  - Flexibility in energy profile.
  - Flexibility in time scheduling.

Flexibility: the amount of energy and the duration of time to which the device energy profile (energy flexibility) and/or activation time (time flexibility) can be changed."

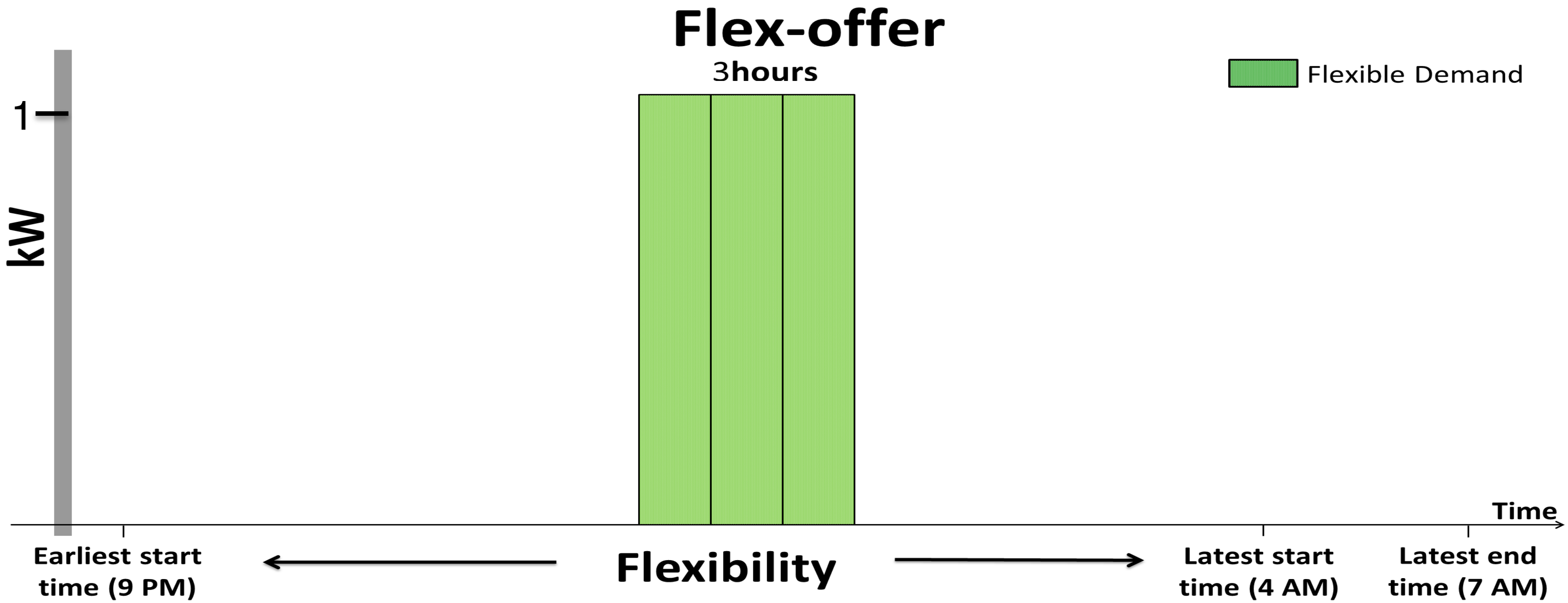
Background



Energy demand and supply, before and after demand flexibility management (using flex-offer).

Background

- TotalFlex[1] project, implements a mechanism to express and utilize the notion of flexibility, using the concept of flex-offer[2].



[1]Totalflex Project, www.totalflex.dk/Forside/
[2]flex-offer, proposed in EU FP7 project MIRABEL, www.mirabel-project.eu

Device Operation Properties

- 1) There exists detectable Intra-day and Inter-day patterns in device operation.
  - (a) Weekend and Weekdays patterns are different.
  - (b) Houses exhibit general and specific intra-day and inter-day patterns.
- 2. There exist time and energy flexibility in device operation.
  - (a) A major percentage of energy consumption comes from flexible devices.
  - (b) An alteration in device energy profile is feasible.
  - (c) Device activation time can be shifted by some duration.
- 3. Some devices are correlated
  - (a) Highly correlated device are operated simultaneously or just after one another.
  - (b) There is some fixed sequence of device operation.

Dataset

- REDD[4] dataset
  - April to June, 2011.

House Number	Days Span	#Days	#Channels	#Devices
House 1	36	35	18	11
House 2	34	15	9	9
House 3	44	23	20	13
House 4	48	30	18	12
House 5	44	9	24	15
House 6	23	18	15	11

Data details for each house.

Device Categorization

- Evaluate devices based on the cost and benefit of utilizing it under the TotalFlex scenario.

Cost: The loss of user-perceived quality caused by accepting flexibility. (for consumers)

Benefit: The available time and energy flexibility for the device.(for energy supplier)

Device Categorization

- Categorization of devices in to three different *flex-categories*
  - Fully-flexible : High benefit at low cost
  - Semi-flexible : Benefit and cost are comparable
  - Non-flexible : Low benefit and high cost

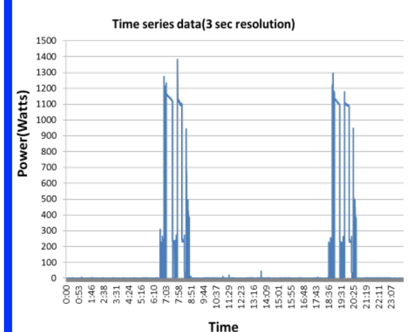
Fully-Flexible	Semi-Flexible	Non-flexible
Dishwasher	Furnace	Bathroom_gfi
Electric_heat	Microwave	Miscellaneous
Refrigerator	Stove	Electronics
Washer_dryer	Oven	Kitchen_outlets
		Lighting

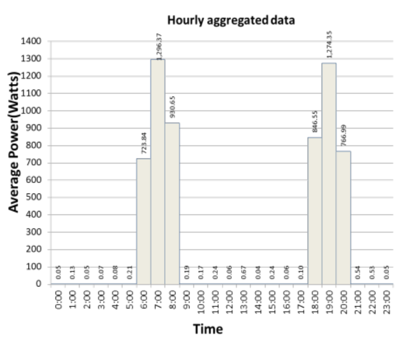
Device flex-categorization

Preprocessing

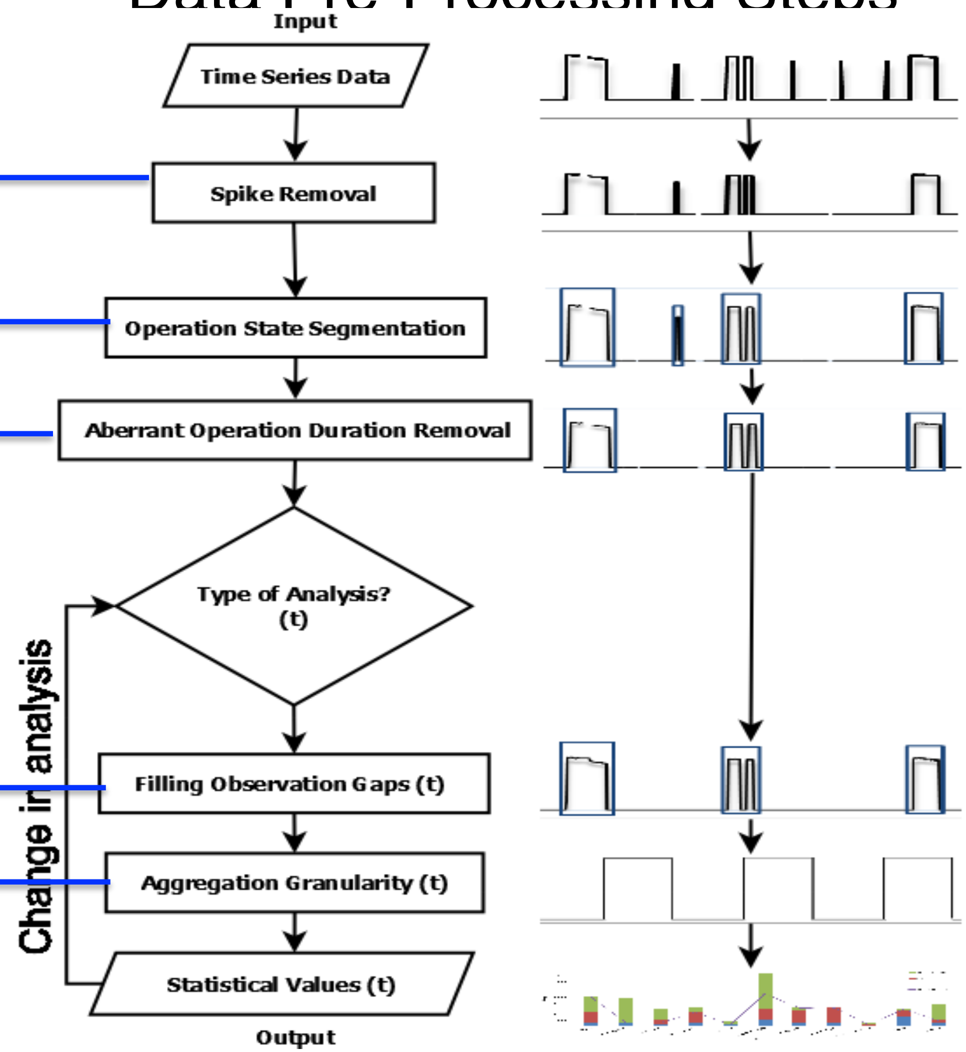
Aggregation Granularity

- Aggregate data into the time granularity that we target for analyses, e.g. hourly or daily.

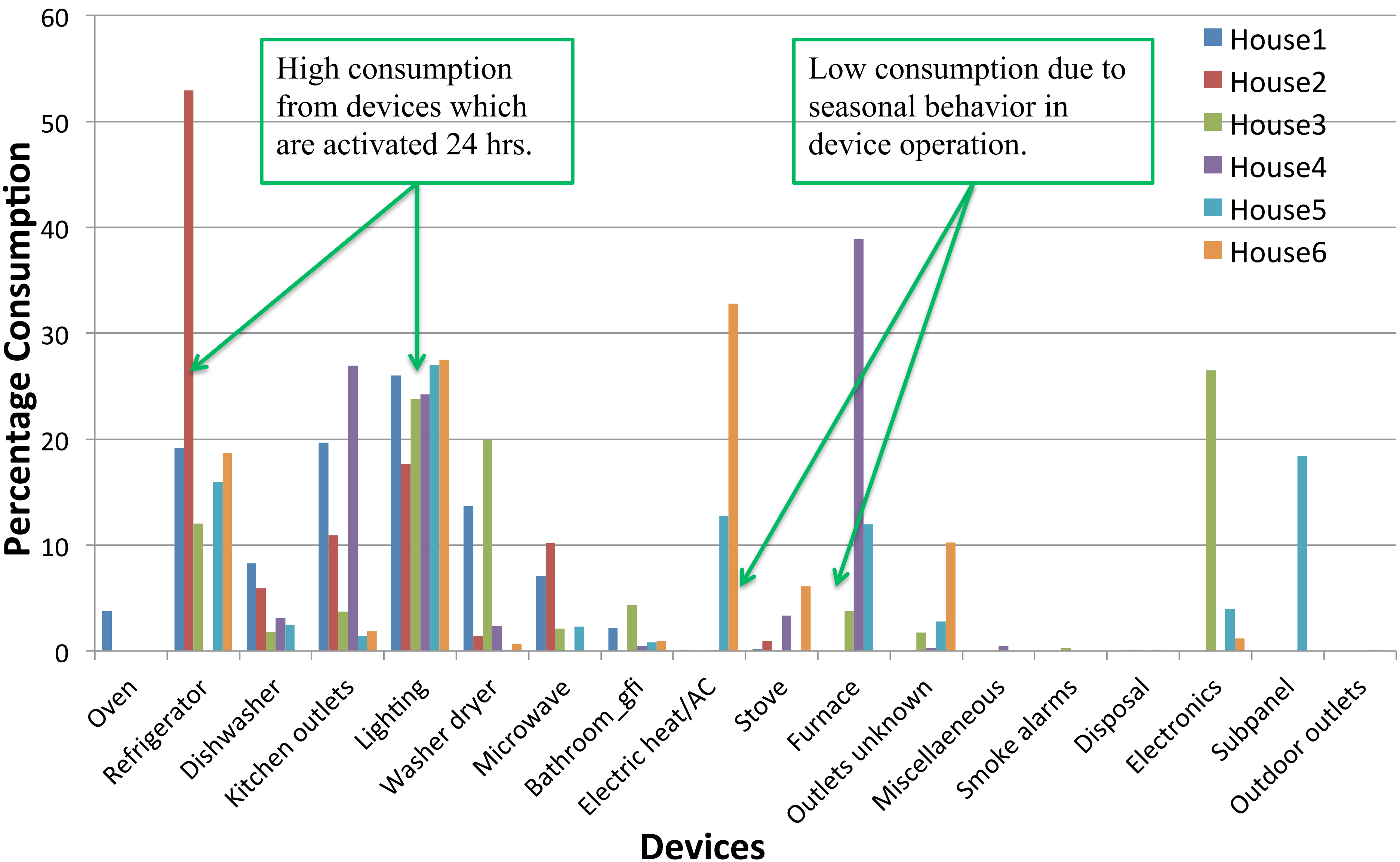




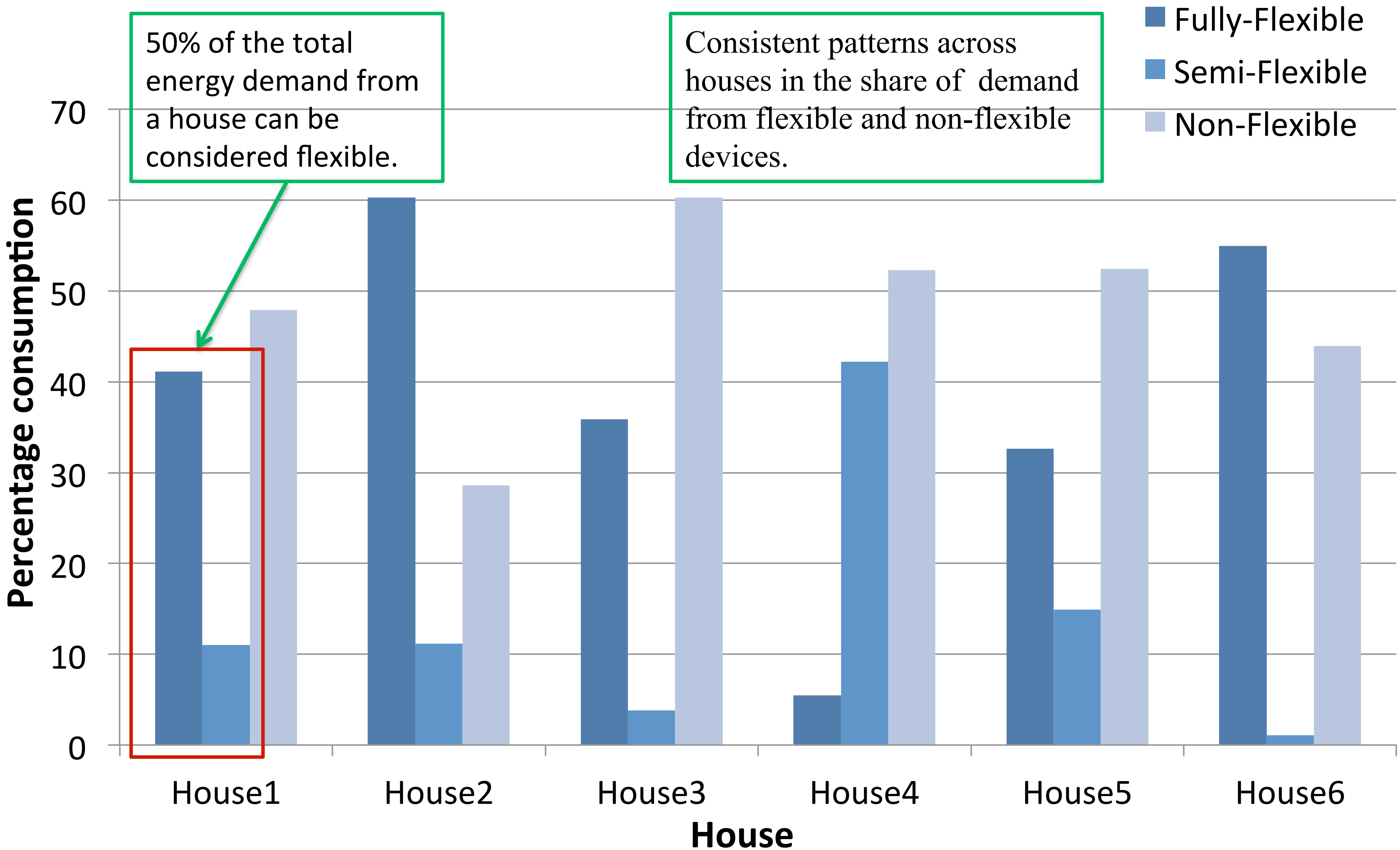
Data Pre-Processing Steps



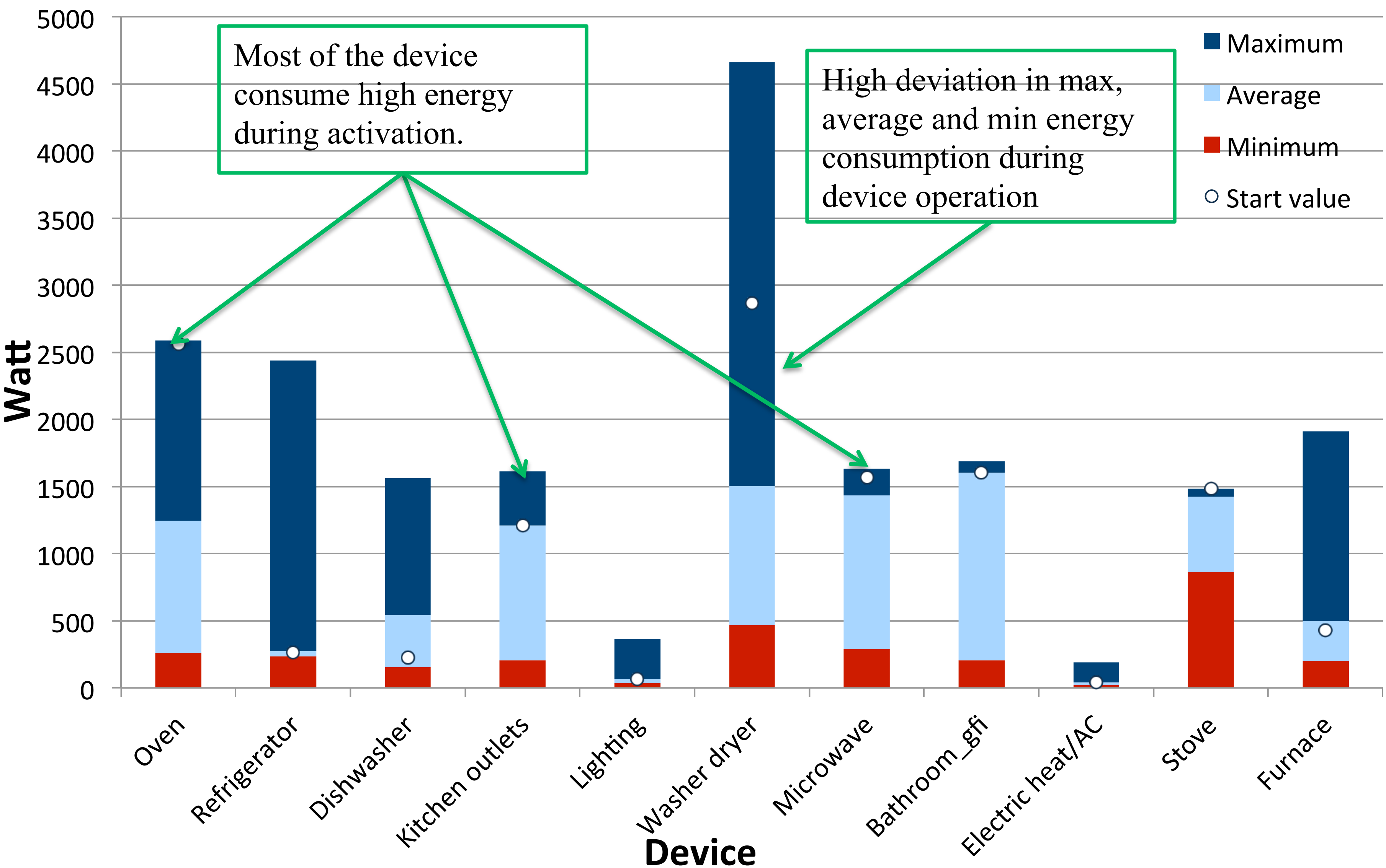
Distribution over various devices



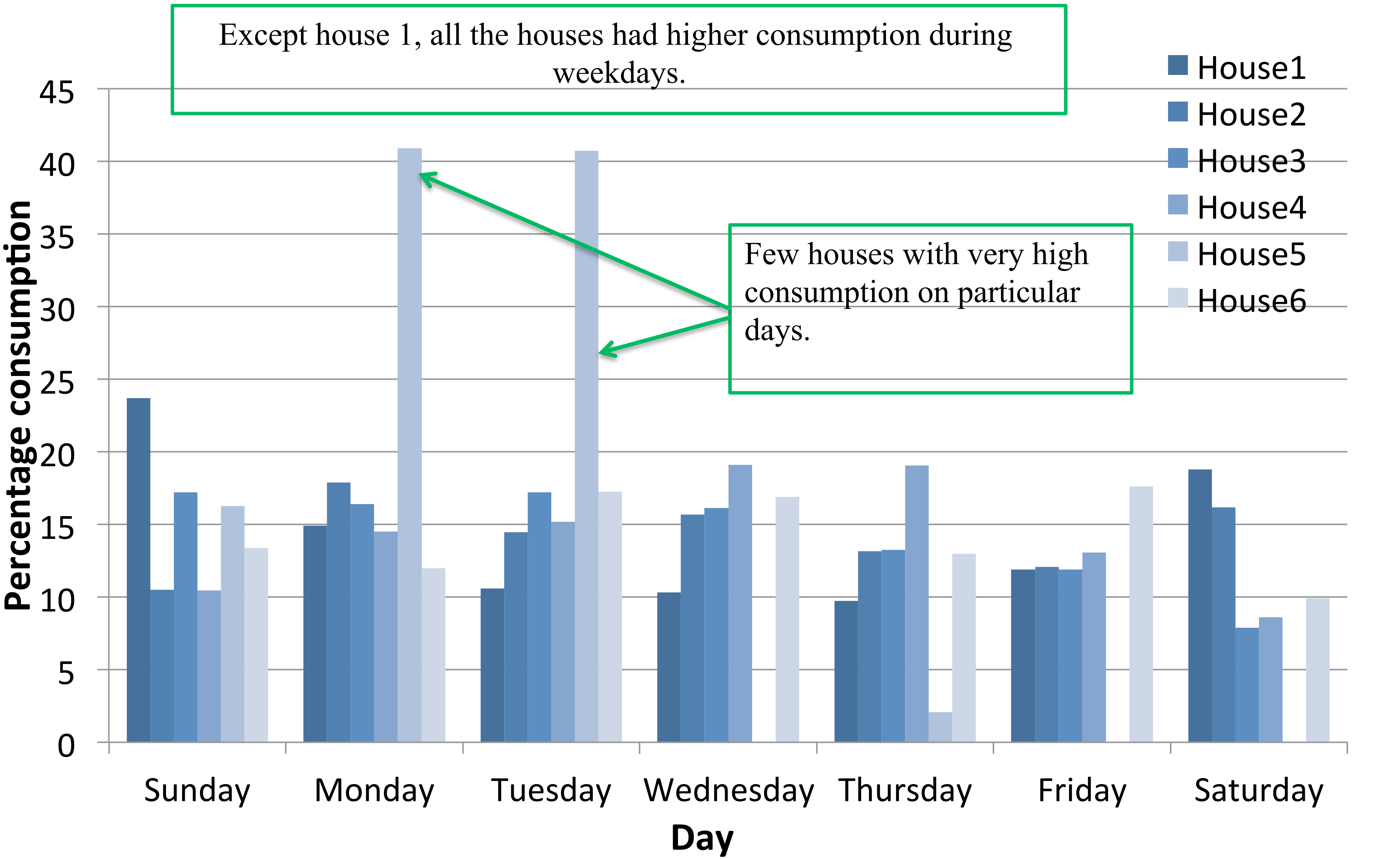
Distribution Over Flexibility Types



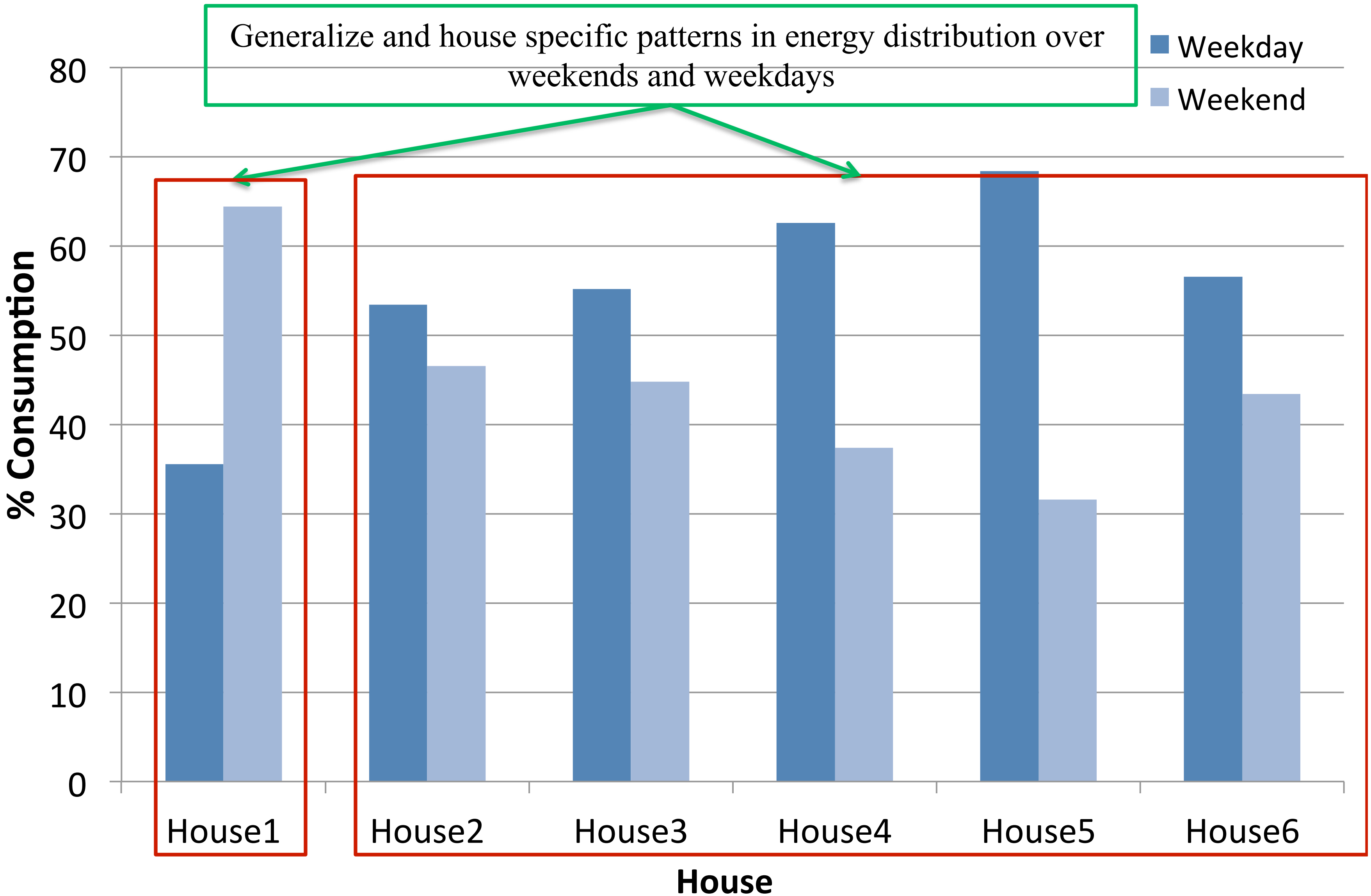
Min, Avg, and Max power Consumption



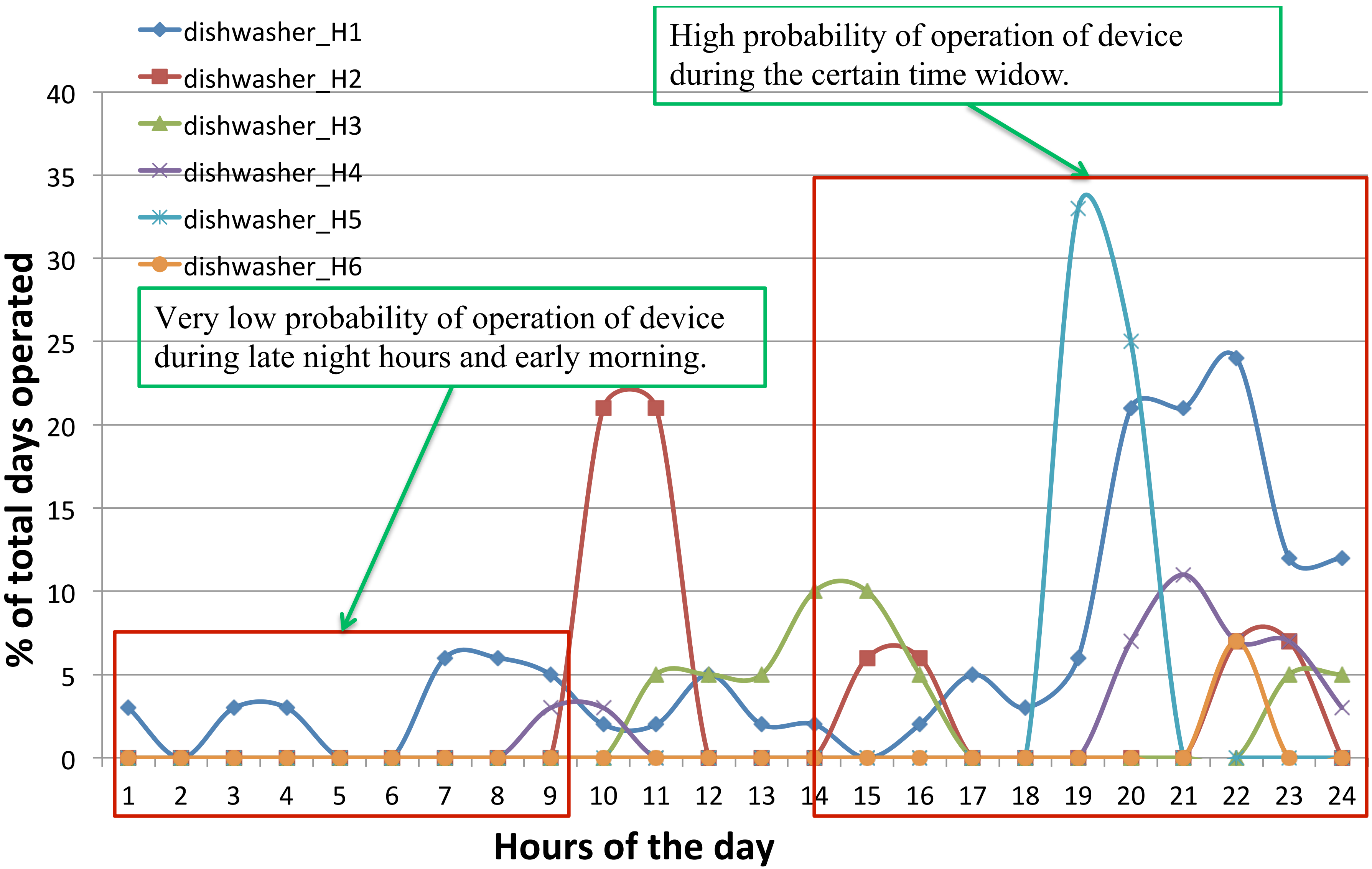
Distribution Over Days



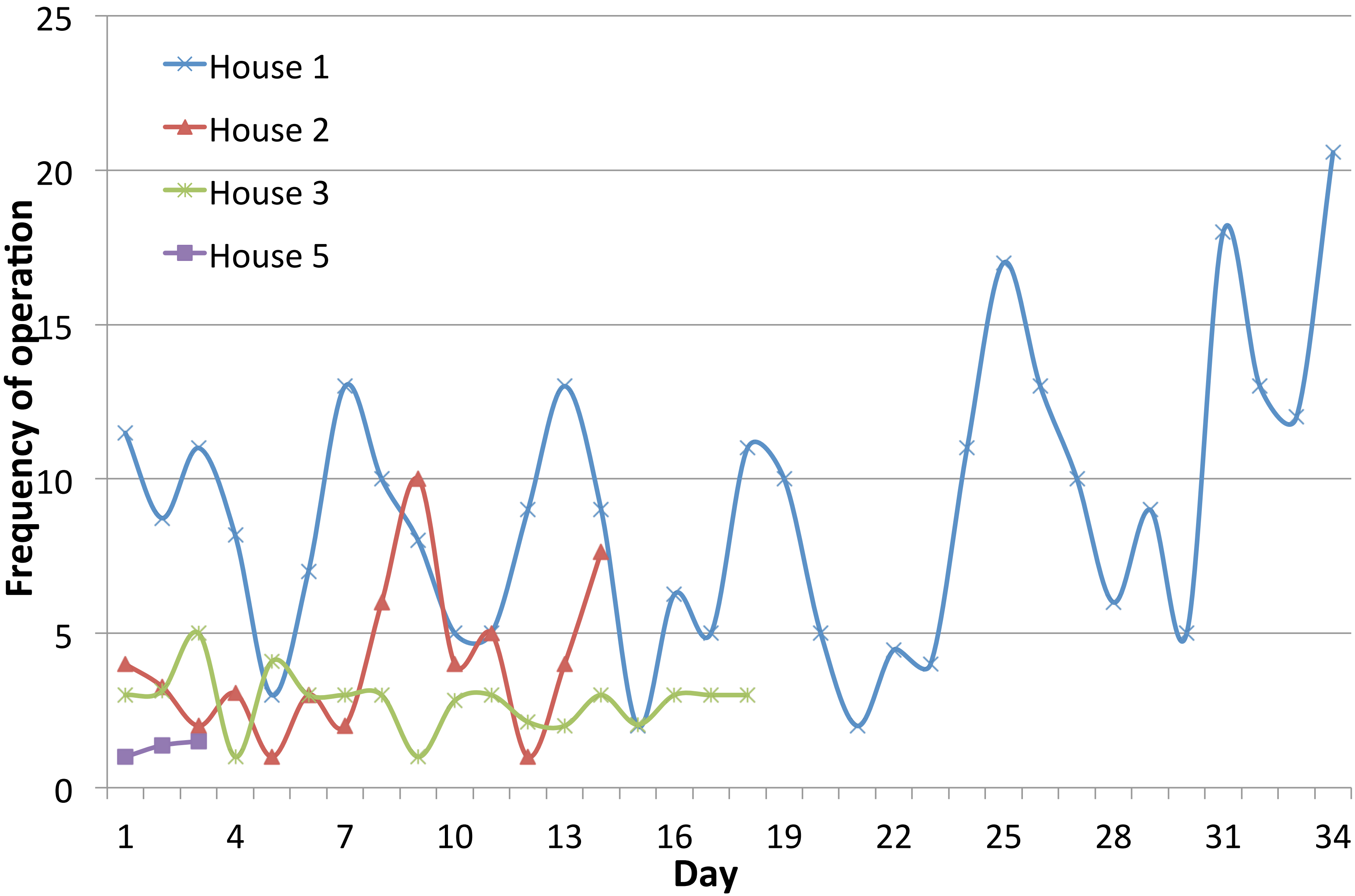
Weekdays Vs Weekends Distribution



Distribution of Hourly Device Operations



Daily Operation Frequency



Device Correlations

Device1	Device2	Frequency	Device1	Device2	Frequency
Oven	Washer dryer	2	Washer dryer	Oven	2
Oven	Microwave	8	Microwave	Oven	9
Oven	Electric heat	1	Electric heat	Oven	1
Dishwasher	Oven	1	Oven	Dishwasher	0
Dishwasher	Washer dryer	2	Washer dryer	Dishwaser	3
Dishwasher	Microwave	2	Microwave	Dishwasher	10
Dishwasher	Stove	1	Stove	Dishwasher	0
Washer dryer	Microwave	12	Microwave	Washer dryer	10
Washer dryer	Electric heat	1	Electric heat	Washer dryer	1
Microwave	Electric heat	8	Electric heat	Microwave	4
Microwave	Stove	6	Stove	Microwave	2
Electric heat	Stove	4	Stove	Electric heat	2
Stove	Oven	1	Oven	Stove	0

Operation sequence for pairs of devices (house 1).

Operation Properties Revisited

- 1) There exists detectable Intra-day and Inter-day patterns in device operation.
  - (a) Weekend and Weekdays patterns are different. ✓
  - (b) Houses exhibit general and specific intra-day and inter-day patterns. ✓
- 2. There exist time and energy flexibility in device operation.
  - (a) A major percentage of energy consumption comes from flexible devices. ✓
  - (b) An alteration in device energy profile is feasible. ✓
  - (c) Device activation time can be shifted by some duration. ✓
- 3. Some devices are correlated
  - (a) Highly correlated devices are operated simultaneously or just after one another ✗
  - (b) There is some fixed sequence of device operation. ✓

Flexibility Study Summary

- Significant percentage of the total energy demand for a house can be considered to provide flexibility.
- Repeating inter-day and intra-day, house-specific or general patterns across houses.
- Potential of extracting time flexibility.
- Potential of extracting energy flexibility.
- There exist interesting correlations and sequences between device operation.
- Patterns and periodicities for device operation can be detected and predicted, even in stochastic environments.

Flexibility Study Conclusion

- User’s possess flexibility in their usage patterns.
- These flexibility can be extracted with low loss of user perceived quality.
- Support the concept of the TotalFlex project of utilizing flexibility for demand management.

Future Work

1. Design models for flexibility- and load prediction.
2. Econometric analysis of flexibility.
3. Generation of flex-offers.

Ongoing Project: Arrowhead

- Collaborative automation
  - Equipment, people, and IT services work together to optimize
  - Largest EU FP7 project
  - FOs as the basis for a Virtual Market of Energy
  - Generic service-oriented architecture for optimal integration
  - Demonstrators/trials for residential buildings, commercial buildings, industrial processes, electromobility (EVs)
  - www.arrowhead.eu

Ongoing and Future Work

- **Constraint** aggregation
  - Aggregate flex-offers so that they respect **grid constraints**
- Demand forecasting at **device level**
  - Challenge of stochastic behavior
- Flexibility **detection**
  - Extracted from device level forecasts
  - Challenge to estimate available flexibility
- Flexibility **prediction**
  - What flexibility will be available tomorrow
  - Learn the behavior of users and their flexible devices
- Flex-offer generation
  - Based on predicted flexibilities
- Markets and tax schemes for flexibility
- Integration in devices and systems of systems

Big Energy Data Summary

- Why?
  - CO_2 reductions, more renewable energy sources
  - Make (flexible) demand meet (renewable) supply
- What is it?
  - Time series of demand and supply
  - **Flex-offers**: generalized and explicit energy flexibilities
- What do we do with it?
  - (Repeated) Forecasting, scheduling, ..
  - Storage and querying in a DW
  - Aggregation (incremental, balance)
  - Flexibility detection and extraction
- Bottom line
  - Many data management challenges
  - Some domain specific, some general
  - Join the fun ☺

Key References

- http://people.cs.aau.dk/~tbp
- MIRABEL project: http://mirabel-project.eu/
- Totalflex project: http://www.totalflex.dk
- Arrowhead project: http://www.arrowhead.eu
- Energy Data Management workshop series@EDBT http://www.endm.org/
- Böhm et al. *Data management in the MIRABEL smart grid system*. EnDM 2012
- Doms et al. *MIRABEL – Efficiently managing more renewable energy using explicit demand and supply flexibilities*, World Smart Grid Forum 2013 (Best Poster Award)
- Fischer et al. *Real-Time Business Intelligence in the MIRABEL Smart Grid System*. BIRTE 2012
- Khalefa et al. *Model-based Integration of Past & Future in TimeTravel*. PVLDB 5(12), 2012
- Pedersen: Energy Data Management: Where Are We Headed? EDBT/ICDT Workshops'14
- Siksnys et al. *MIRABEL DW: Managing Complex Energy Data in a Smart Grid*. DaWaK'12
- Siksnys et al. *Aggregating and Disaggregating Flexibility Objects*. SSDBM 2012
- Valsomatzis et al. *Balancing Energy Flexibilities Through Aggregation*. DARE 2014
- Neupane et al. *Towards Flexibility Detection in Device-Level Energy Consumption*. DARE 2014

Acknowledgements

- Many slides borrowed from MIRABEL and Totalflex project partners, colleagues, and collaborators