Data Management Platform

Online Advertisement System

Paolo D'Apice

iZENEsoft (Shanghai) Co., Ltd.

November 27, 2013

Data Management Platform

1. manages audience data for advertisement
   - first-party data
   - third-party data
2. defines new target segments by
   - merging first party data
   - optimizing marketing budget

Data Management Platform

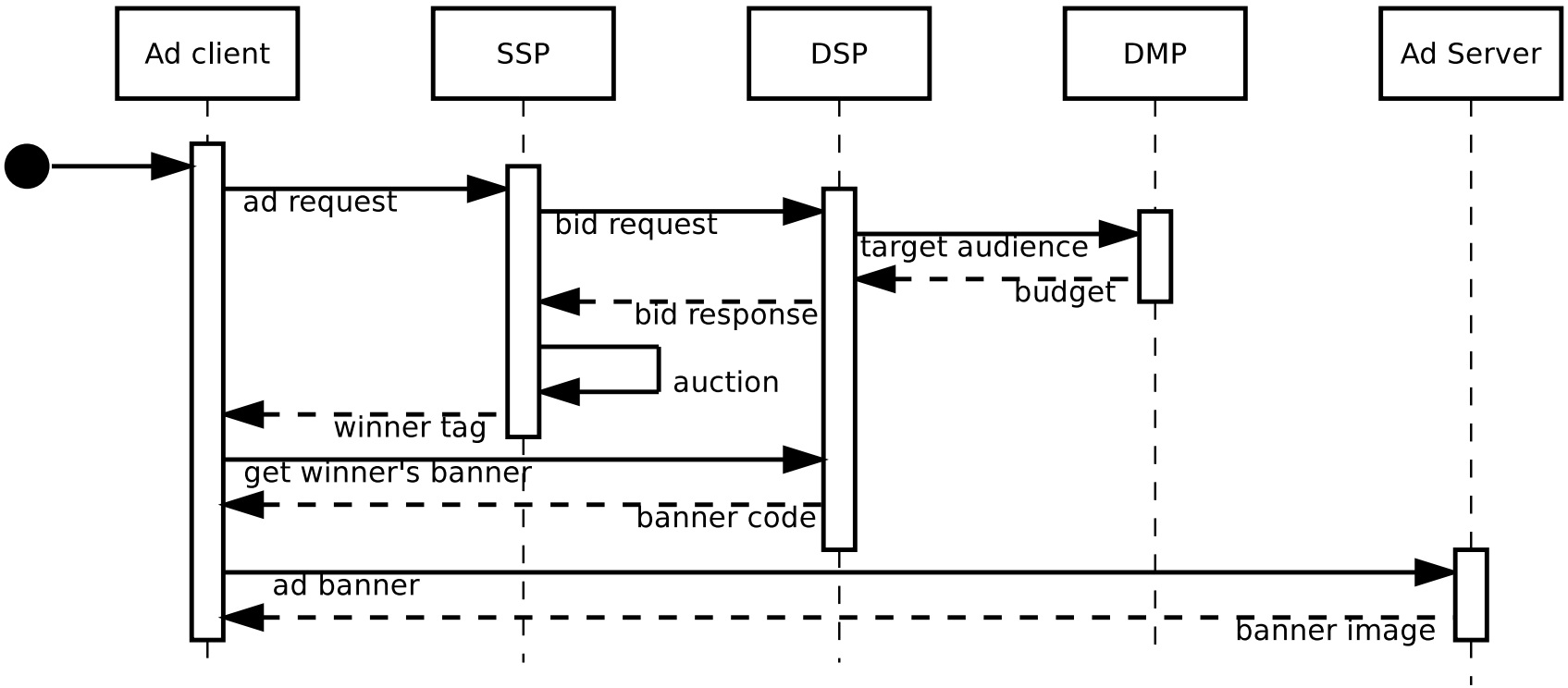
first party data

- behavior, actions or interests across the company's website
- data in CRM
- subscription data
- social data
- cross-platform data

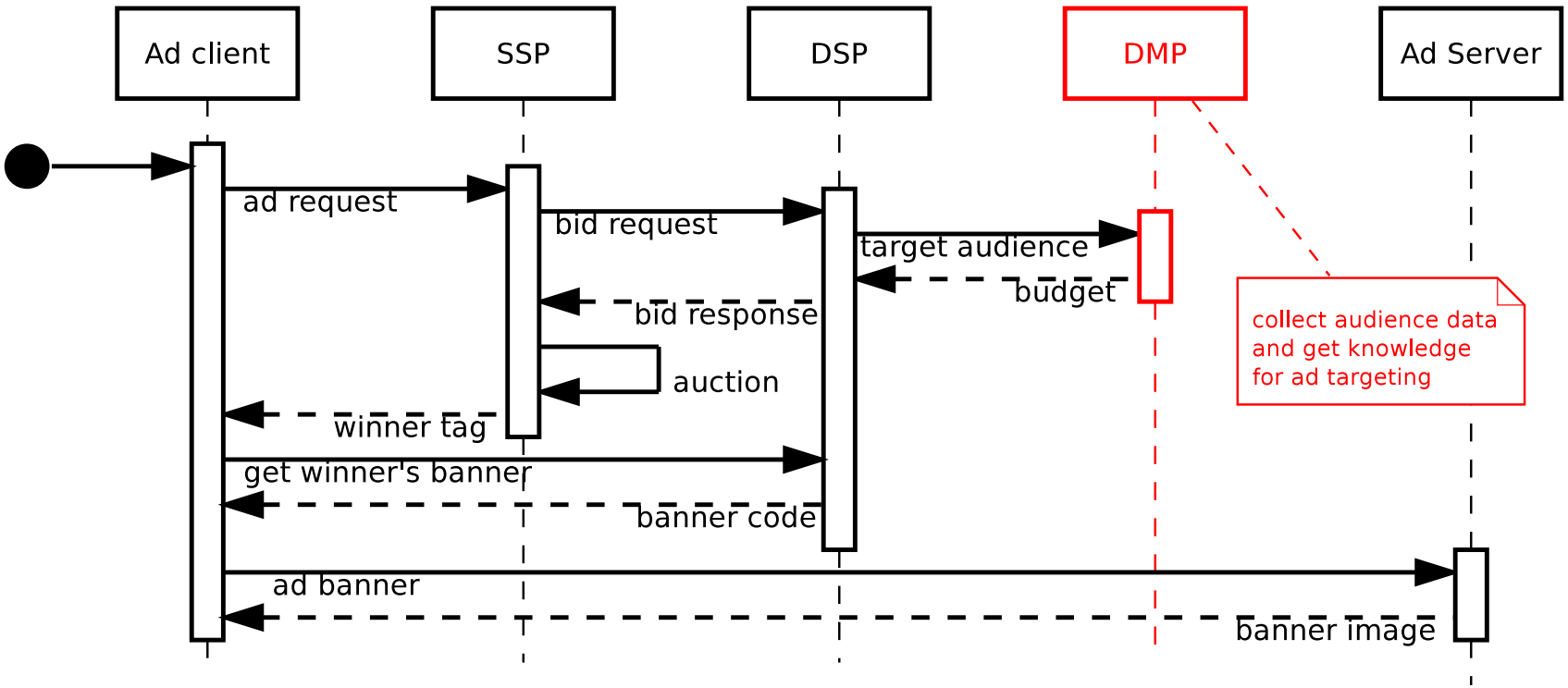
third party data

- generated on other platforms
- aggregated from other websites
- provided by other companies

Data Management Platform



Data Management Platform



DMP and DSP

In-house

Audience data

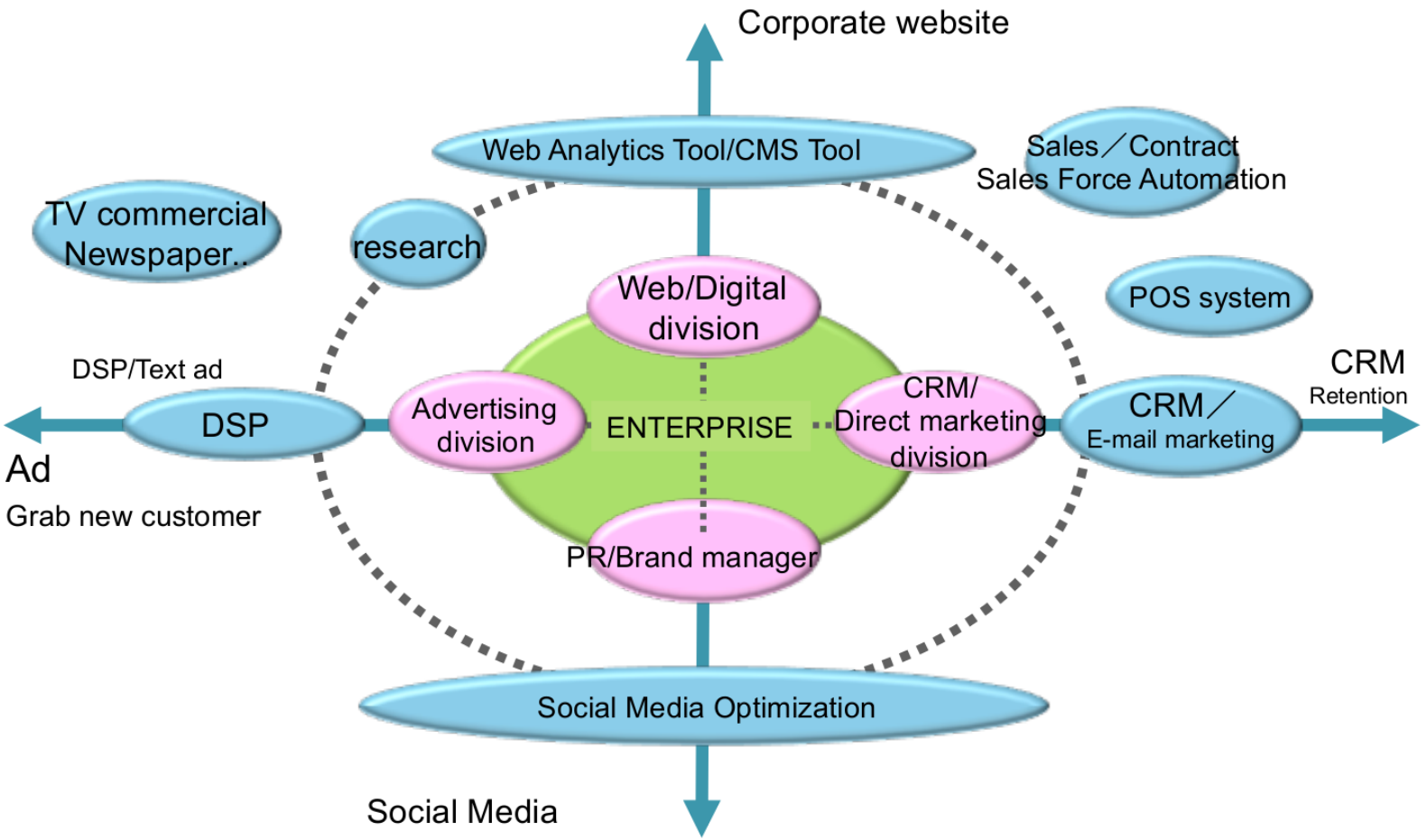
- collect *first party* audience data
- merge with third party data

Advertisement

- define *new* market segments
- optimize marketing budget

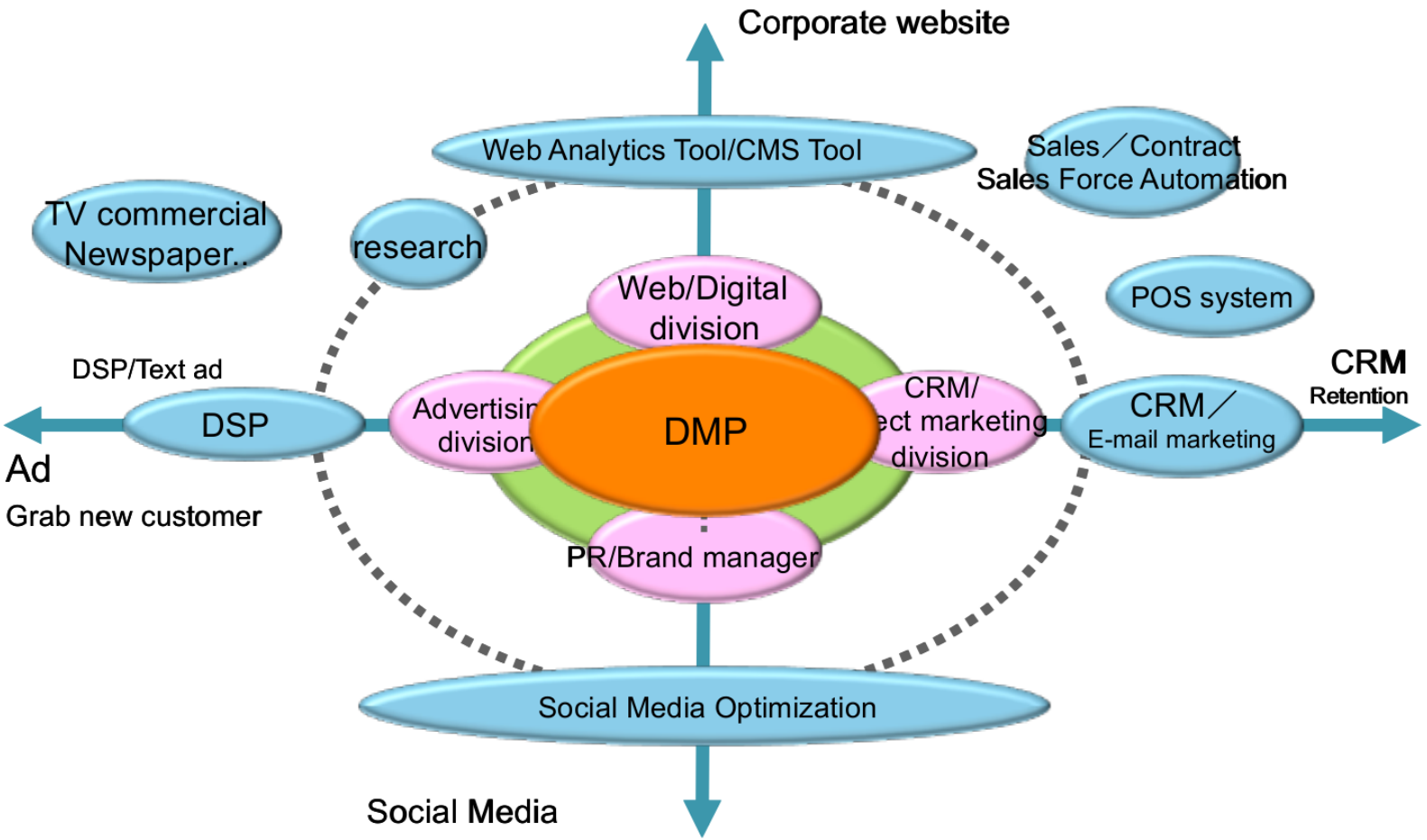
DMP and DSP

How to integrate?



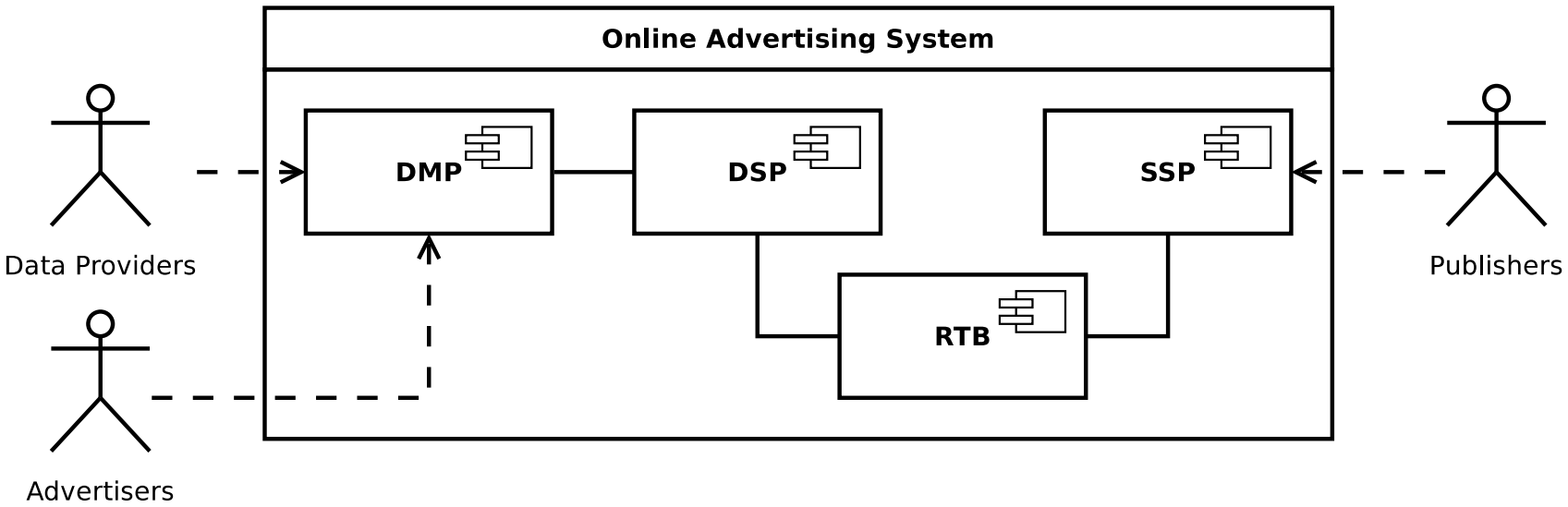
DMP and DSP

How to integrate?



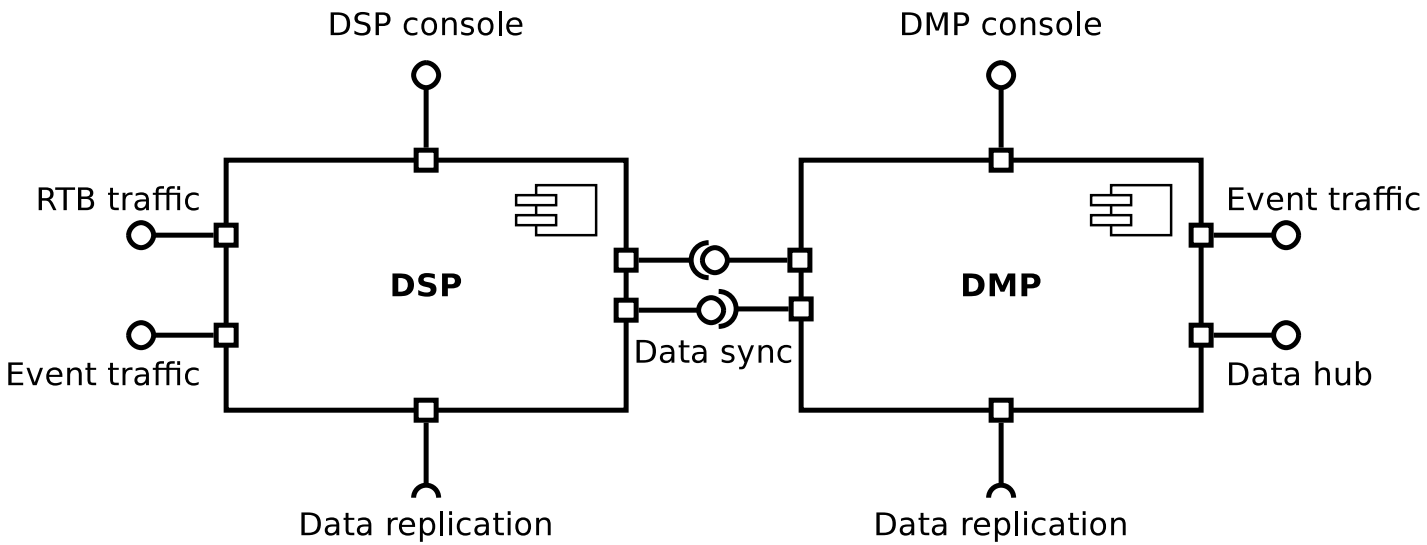
DMP and DSP

Positioning in the Online Advertisement Ecosystem



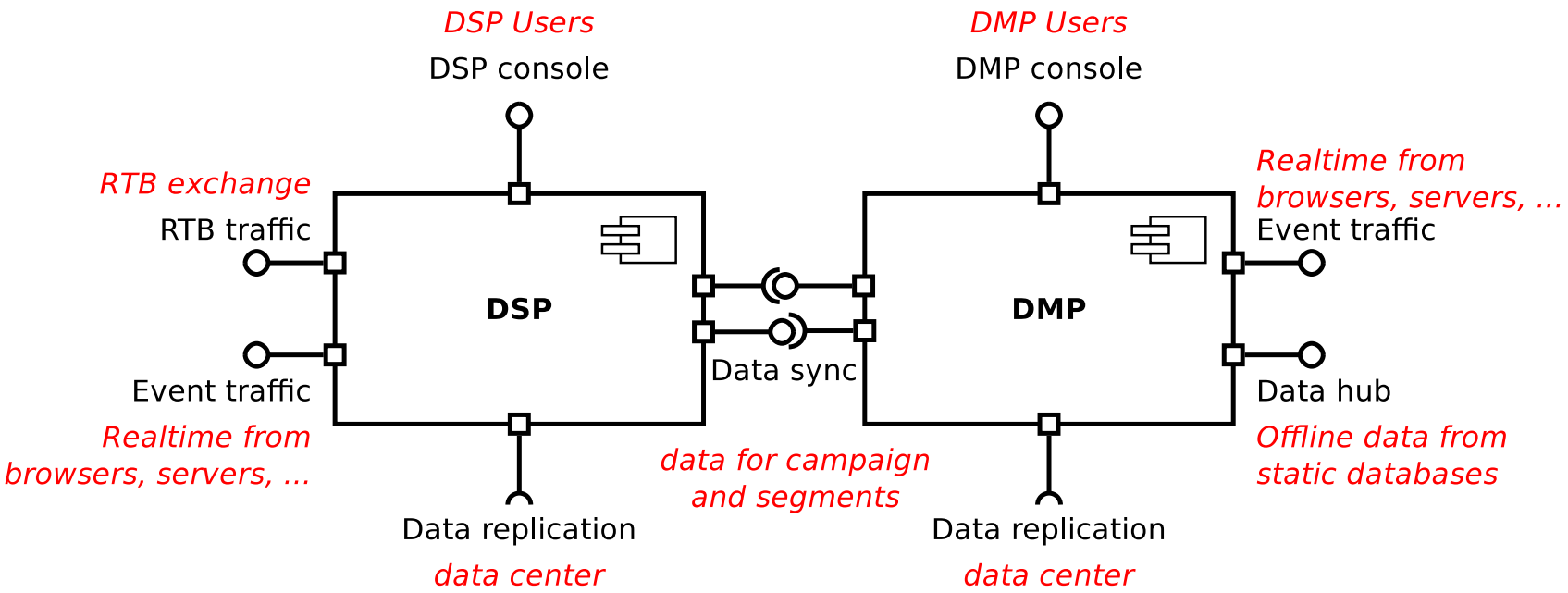
DMP and DSP

Interaction



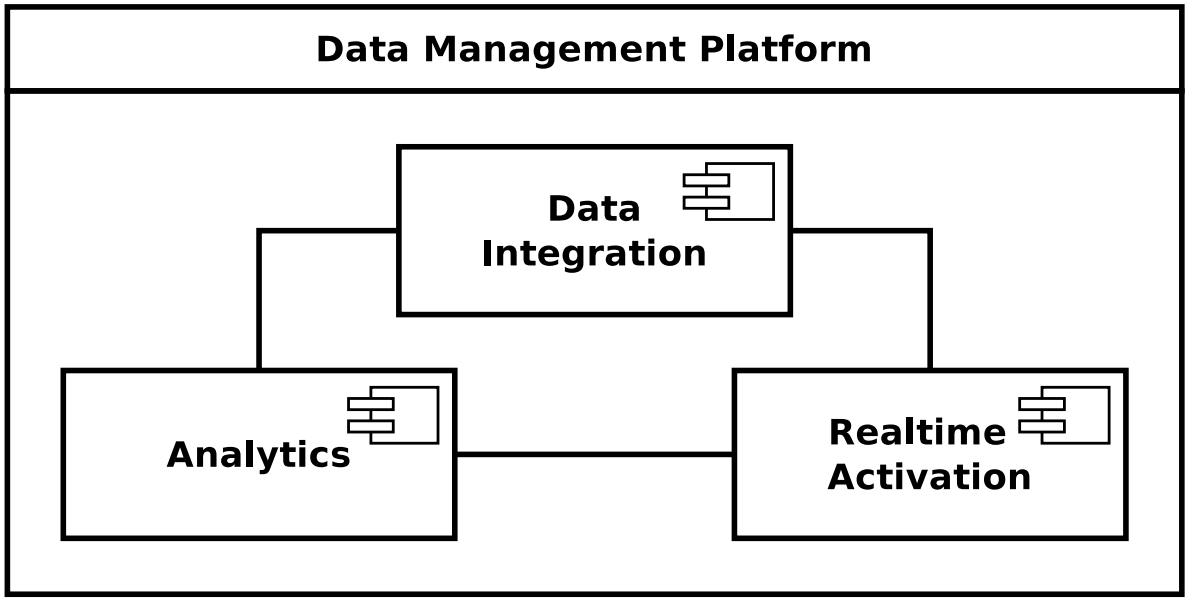
DMP and DSP

Interaction



DMP

Components



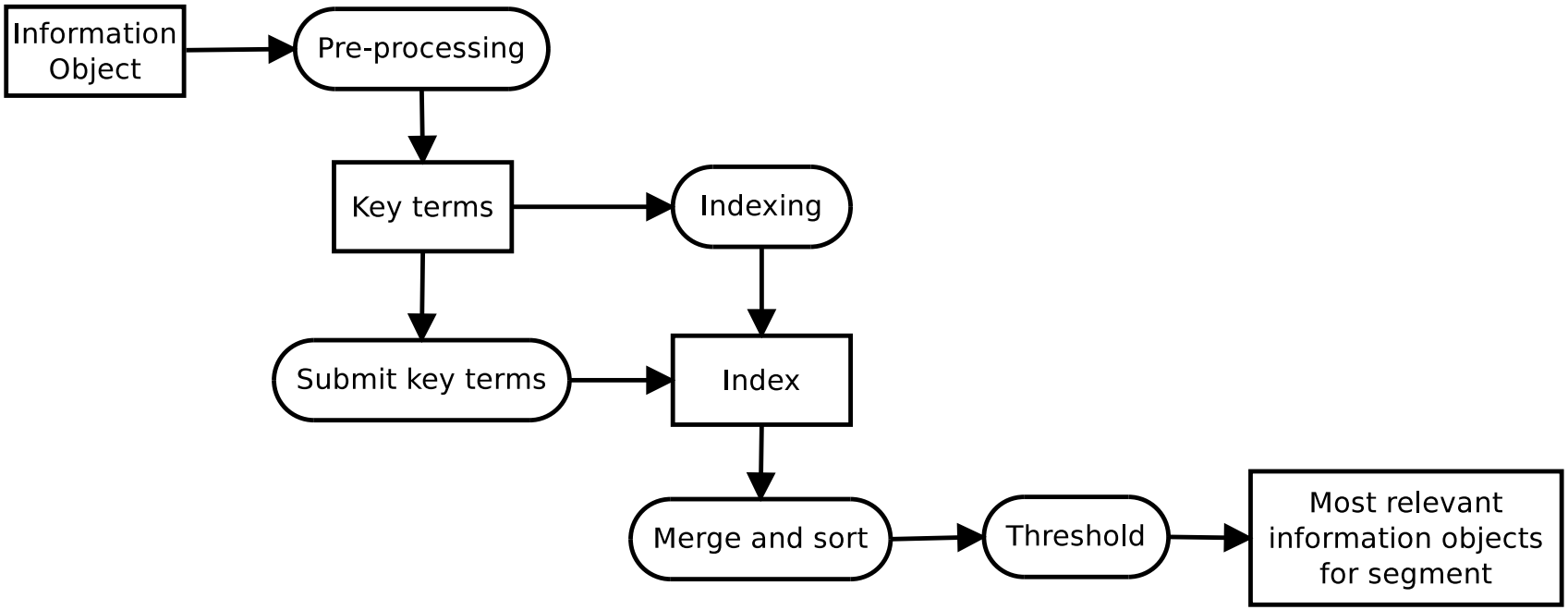
1. Data Integration
   - Schema integration
   - Record linkage
   - Data sanitation
2. Analytics
   - Query definition and optimization
   - Machine Learning
   - Data Mining
3. Real-time activations
   - RUP fast storage
   - Segment evaluation
   - Online classification

Contextual Analysis and Segmentation

1. Analyze pages and create segments
   - N-gram analysis on page content (text and images)
   - terms scoring and thresholding
   - segment assignment by matching
2. Multi-level segmentation
   - categories (statically defined by humans)
   - subcategories (dynamically created and updated)
3. Information Retrieval
   - term indexing and retrieval
4. Hierarchical segmentation

Contextual Analysis and Segmentation

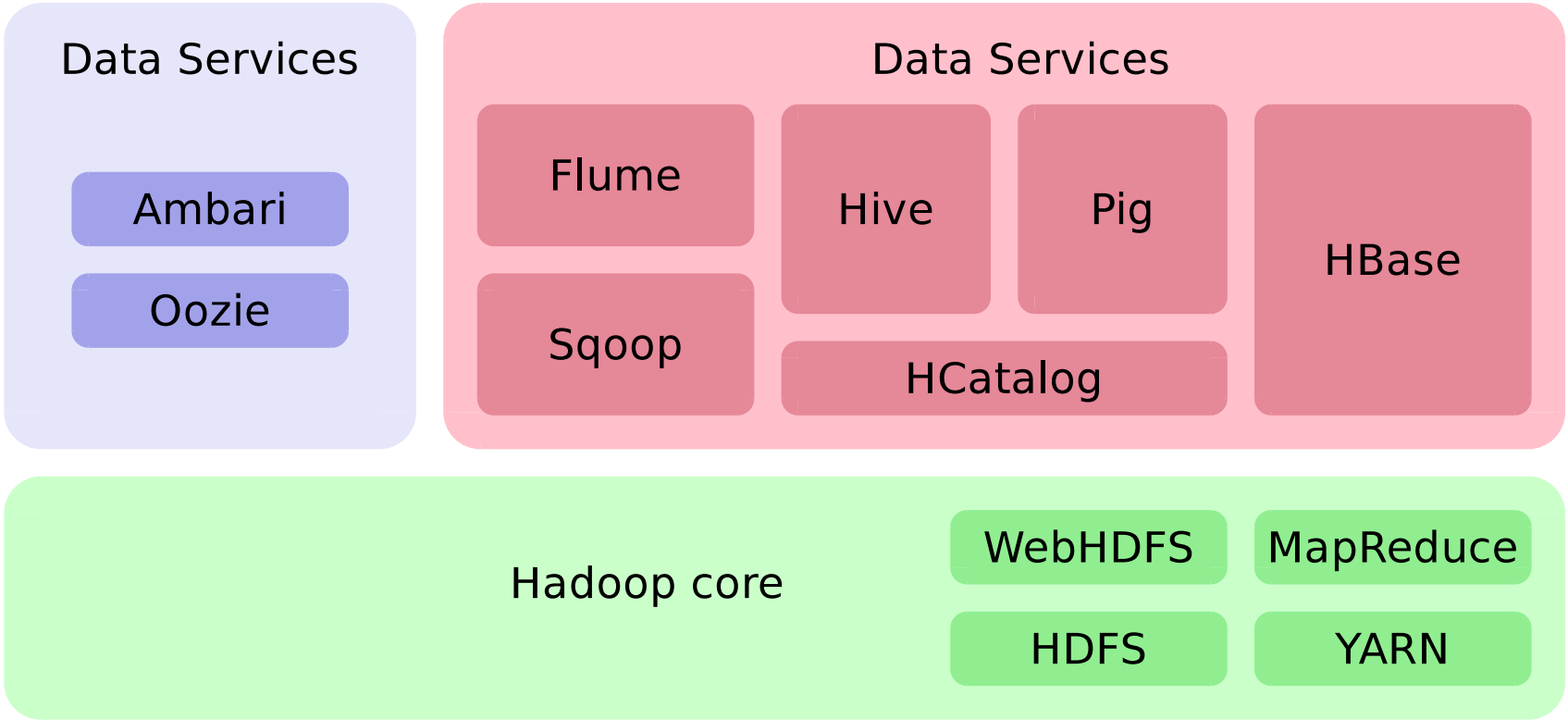
Summary



Behavioral Targeting

- Linear Poisson Regression
  - $p(y) = \frac{\lambda^y e^{-\lambda}}{y!}, \lambda = \mathbf{w}^T \mathbf{x}$
- Estimate CTR as $C\hat{T}R_i = \frac{\lambda_i^{click} + \alpha}{\lambda_i^{view} + \beta}$
  - $\frac{\alpha}{\beta}$ default CTR for new users
- Expected click/views are computed *offline* and updated incrementally in real-time

Hadoop Ecosystem



Data Management Platform

Online Advertisement System

Paolo D'Apice

iZENEsoft (Shanghai) Co., Ltd.

November 27, 2013