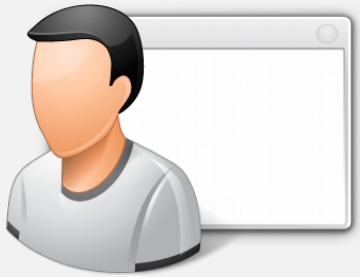
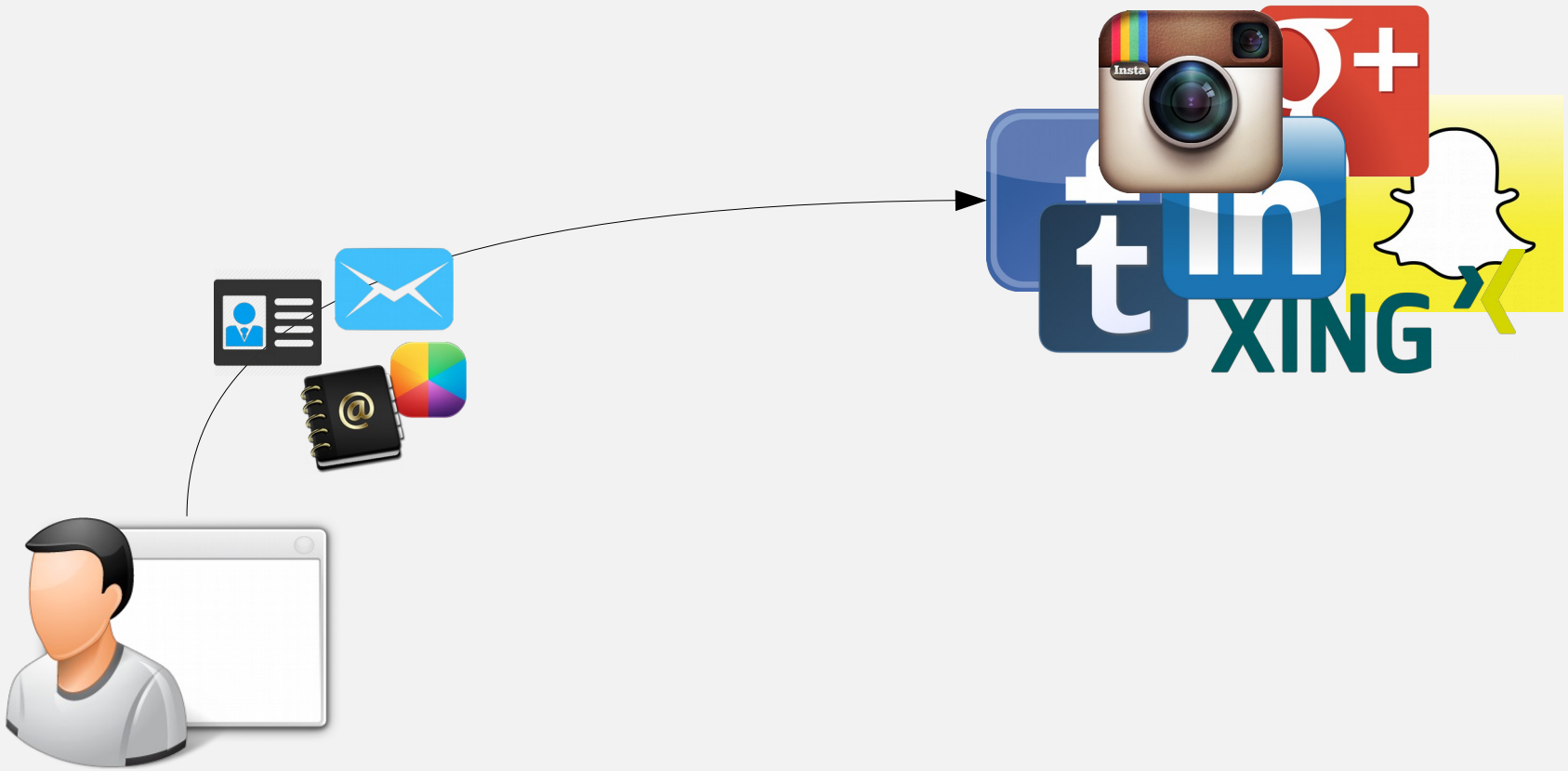
Imperial College London

Compliance Monitoring of Third-Party Applications in Online Social Networks

Florian Kelbert, Imperial College London
Alexander Fromm, Technical University of Munich

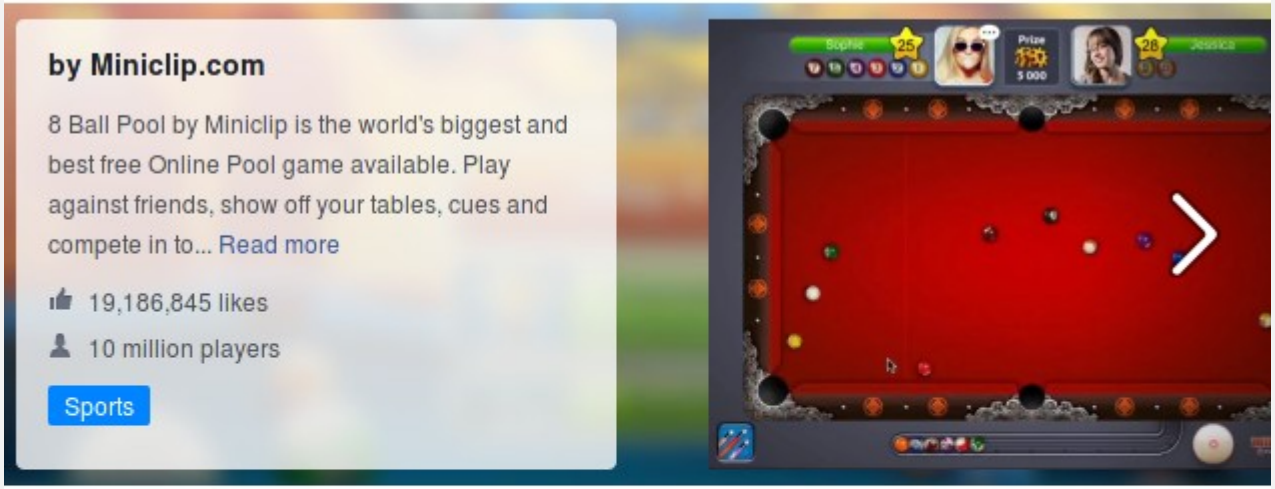



Insta
XING


Insta
XING
Search Facebook
Florian
Home
Find Games
Activity
Search for games...
Home
Top Charts
Casual
Battle
Casino
Platform:
Facebook
Online Soccer Manager (O...
Sports Game by Gamebasics
Show the world that you are the best football manager in this free-to-play football management game.
Top games · See more
9 Ball Pool
500,000 players
Millionaire Quiz
10,000 players
Jelly Match3
100,000 players
darts
50,000 players
Spin It Rich
100,000 players
Recommended Games · See more
Marvel: Avengers Alla
Action
Archery Games
Sports
Snooker Live Pro
Sports
Crash Drive 2
Sports
9 Ball Pool
Sports
8 Ball Pool: Master Tou
Sports
Trending among friends · See more
Candy Crush Saga
Match 3
Casual
Candy Crush Soda Sa
Match 3
Casual
Criminal Case
Puzzle
Casual
Island Experiment
Simulation
Casual
GamePoint Bingo
Bingo
Casino
Klondike
Simulation
Casual

8 Ball Pool

Miniclip.com



Play Now

By clicking on "Play Now" above, **8 Ball Pool** will receive the following info: your public profile, email address and birthday.

Edit the info you provide

By proceeding, you agree to 8 Ball Pool's Terms of Service and Privacy policy

This does not let the app post to Facebook.

Block

Report a Problem

Problem

Problem

How to ensure that data is used in correspondence with policies?

Problem

How to ensure that data is used in correspondence with policies?

Problem

How to ensure that data is used in correspondence with policies?

Problem

How to ensure that data is used in correspondence with policies?

Problem

How to ensure that data is used in correspondence with policies?

Problem

How to ensure that data is used in correspondence with policies?

“You may cache the content for up to 24 hours”

Problem

How to ensure that data is used in correspondence with policies?

“You may cache the content for up to 24 hours”

“Only use friend data in the person’s experience in your app”

Problem

How to ensure that data is used in correspondence with policies?

“You may cache the content for up to 24 hours”

“Only use friend data in the person’s experience in your app”

“You may not disclose confidential information to a third party without the prior explicit consent of Tumblr.”

To start with …

To start with ...

Social Networks are trusted

To start with …

Social Networks are trusted

To start with …

Social Networks are trusted

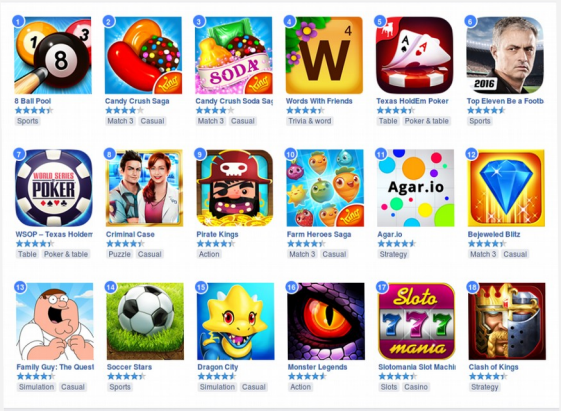
Third Party Applications are not

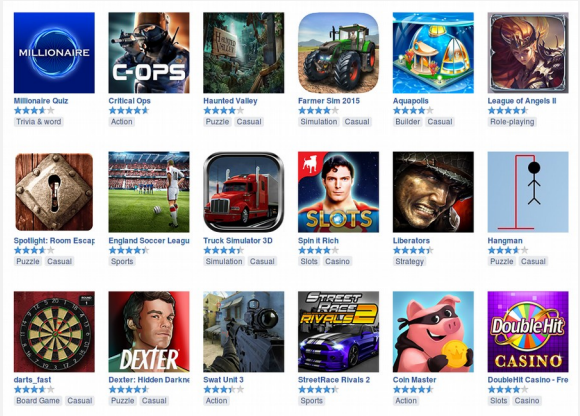
To start with …

Social Networks are trusted

Third Party Applications are not

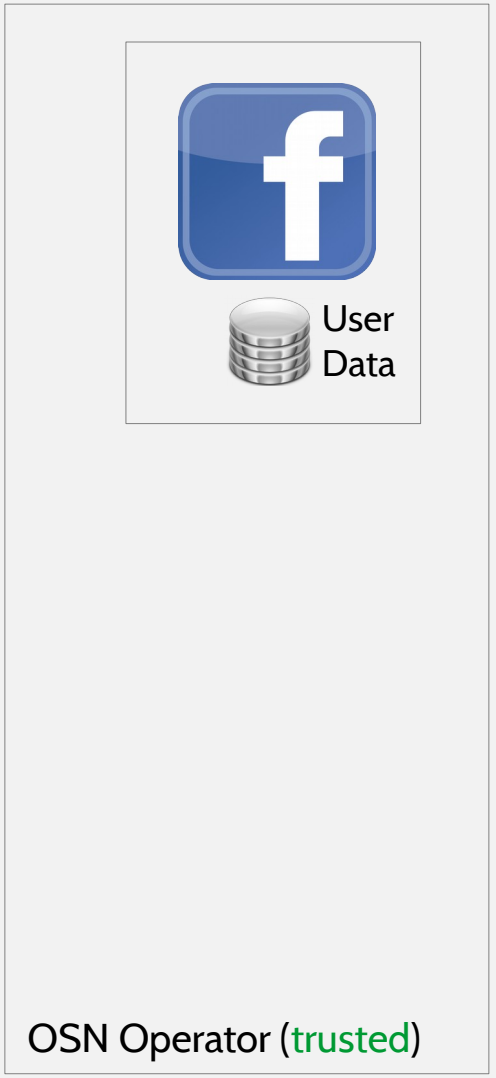
Thousands of apps and developers



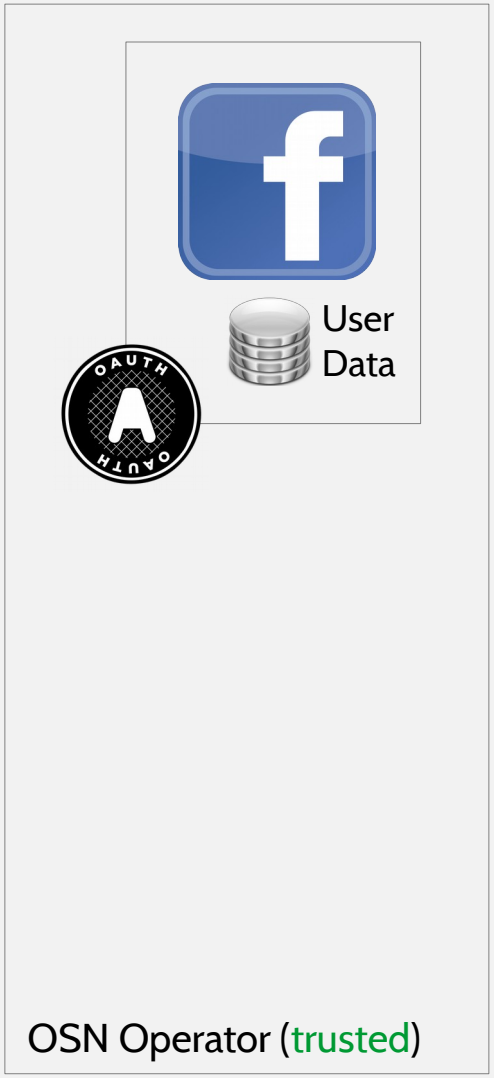


Overview

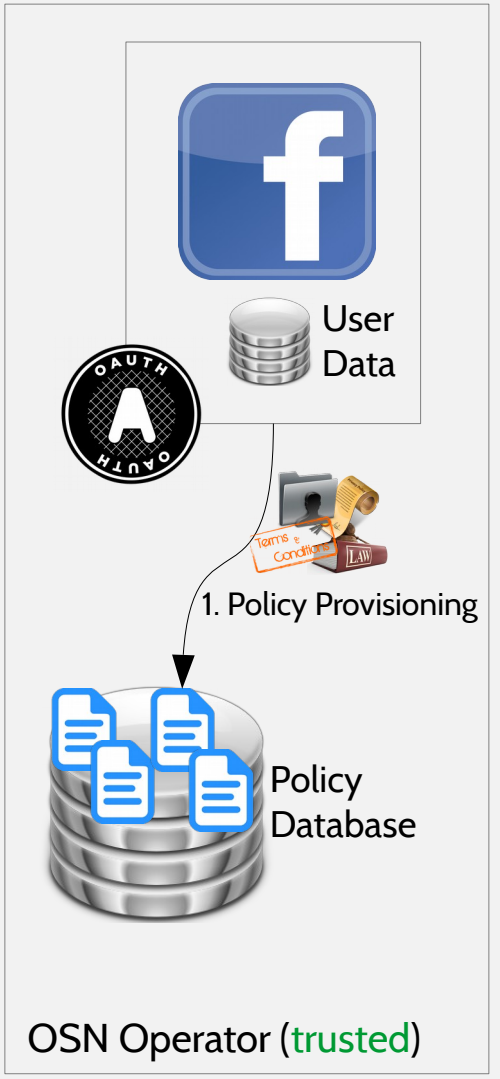
Overview



Overview

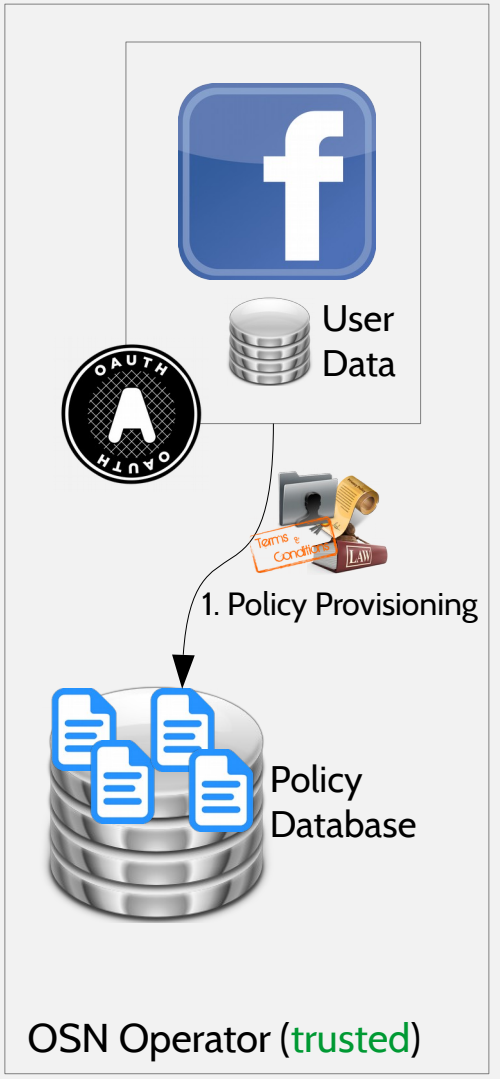


Overview

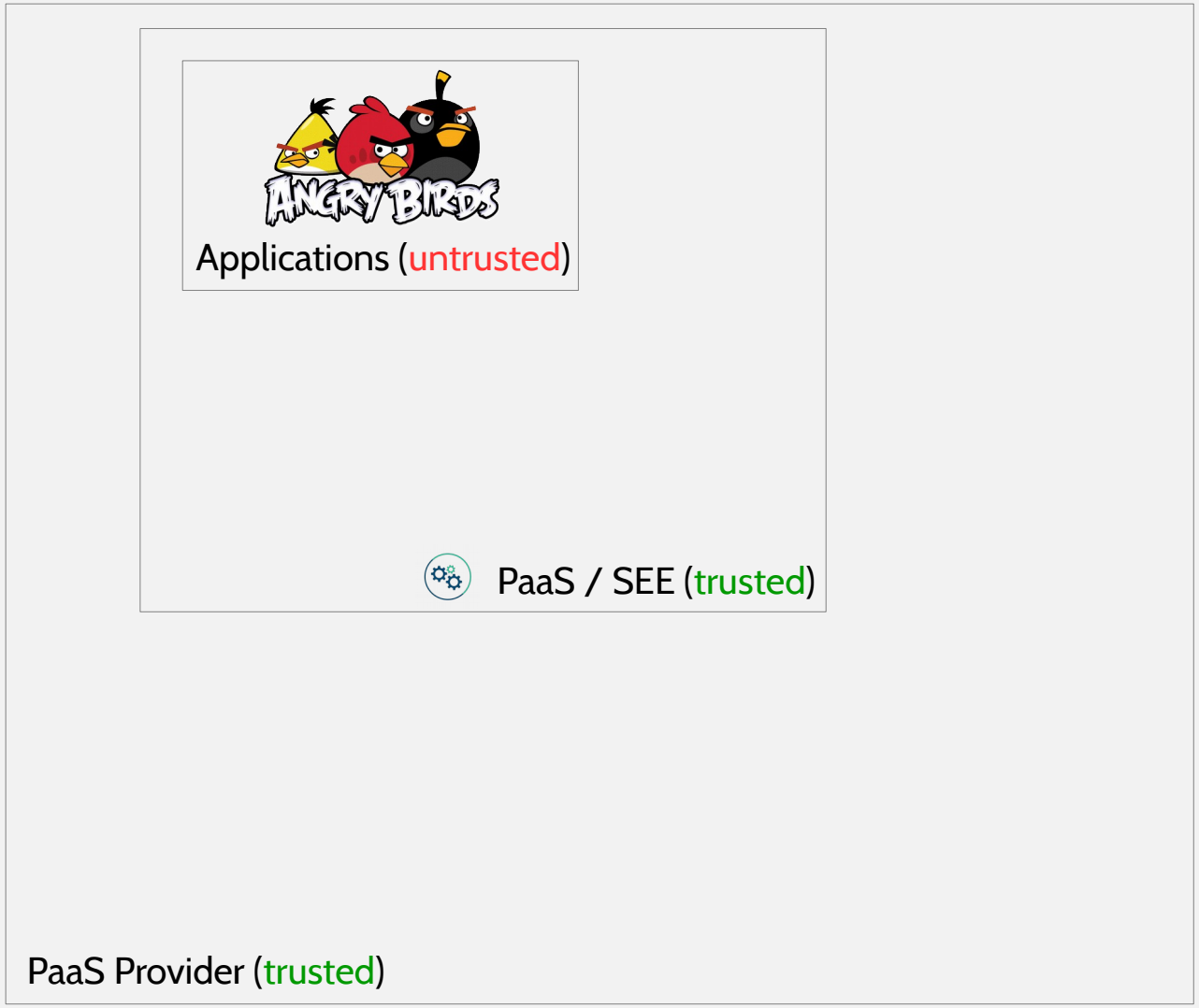

User
Data
1. Policy Provisioning
Policy
Database
OSN Operator (trusted)

Overview


ANGRY BIRDS
Applications (untrusted)


User Data
OAUTH
Terms & Conditions
LAW
1. Policy Provisioning
Policy Database
OSN Operator (trusted)

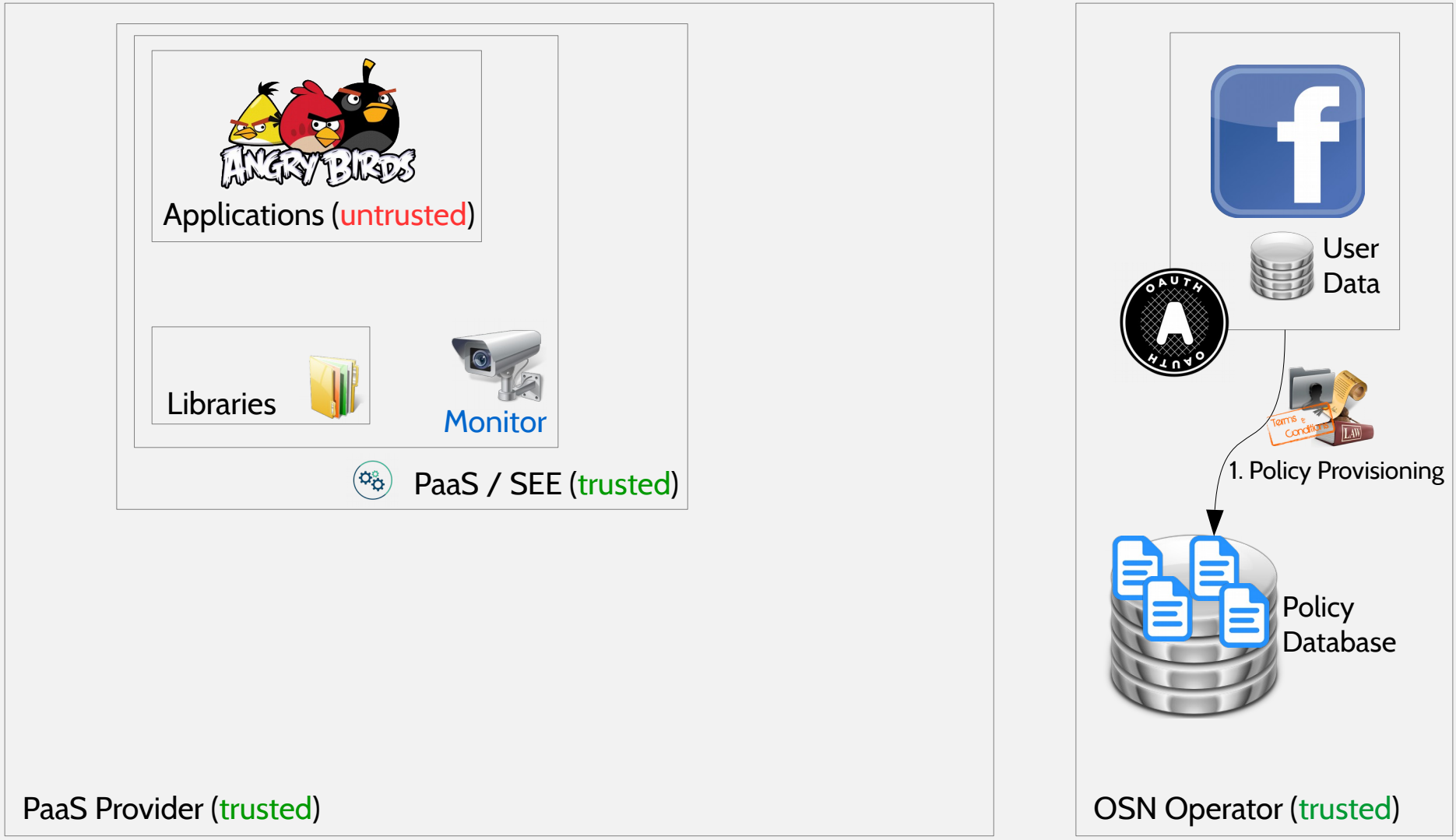
Overview


Applications (untrusted)
PaaS / SEE (trusted)
PaaS Provider (trusted)
User Data
1. Policy Provisioning
Policy Database
OSN Operator (trusted)

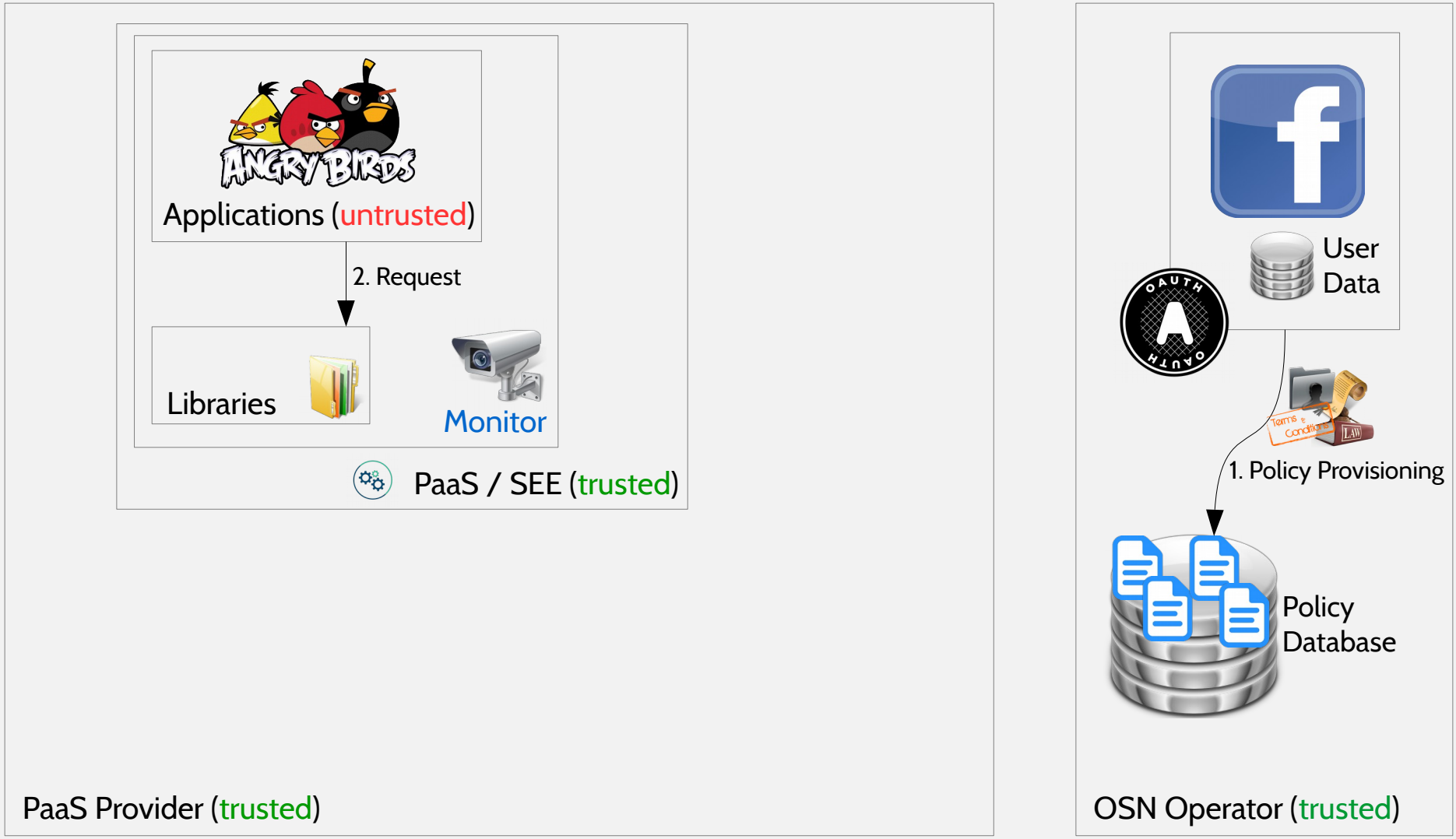
Overview


Applications (untrusted)
Libraries
PaaS / SEE (trusted)
PaaS Provider (trusted)
User Data
1. Policy Provisioning
Policy Database
OSN Operator (trusted)

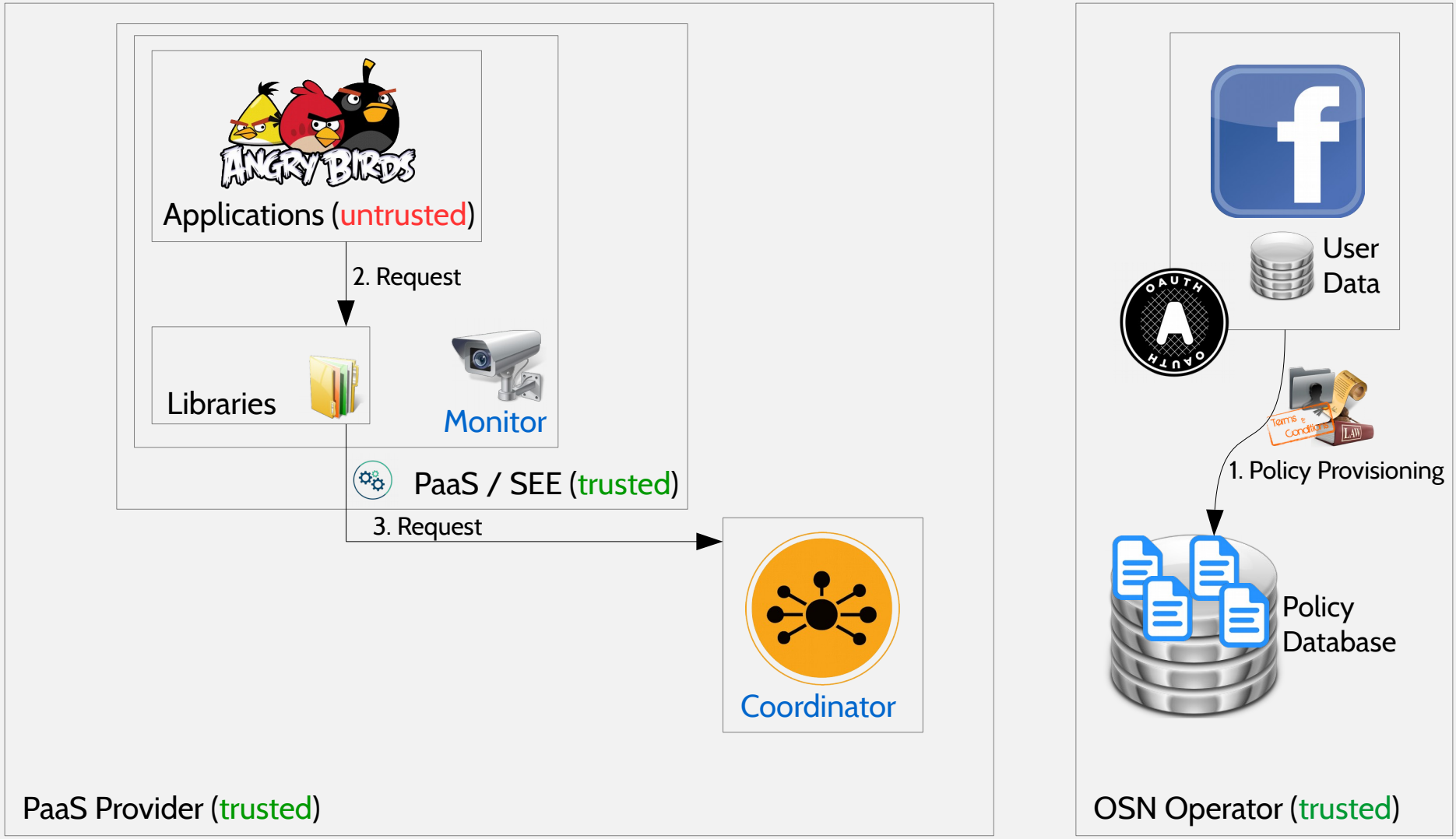
Overview


Applications (untrusted)
Libraries
Monitor
PaaS / SEE (trusted)
PaaS Provider (trusted)
User Data
1. Policy Provisioning
Policy Database
OSN Operator (trusted)

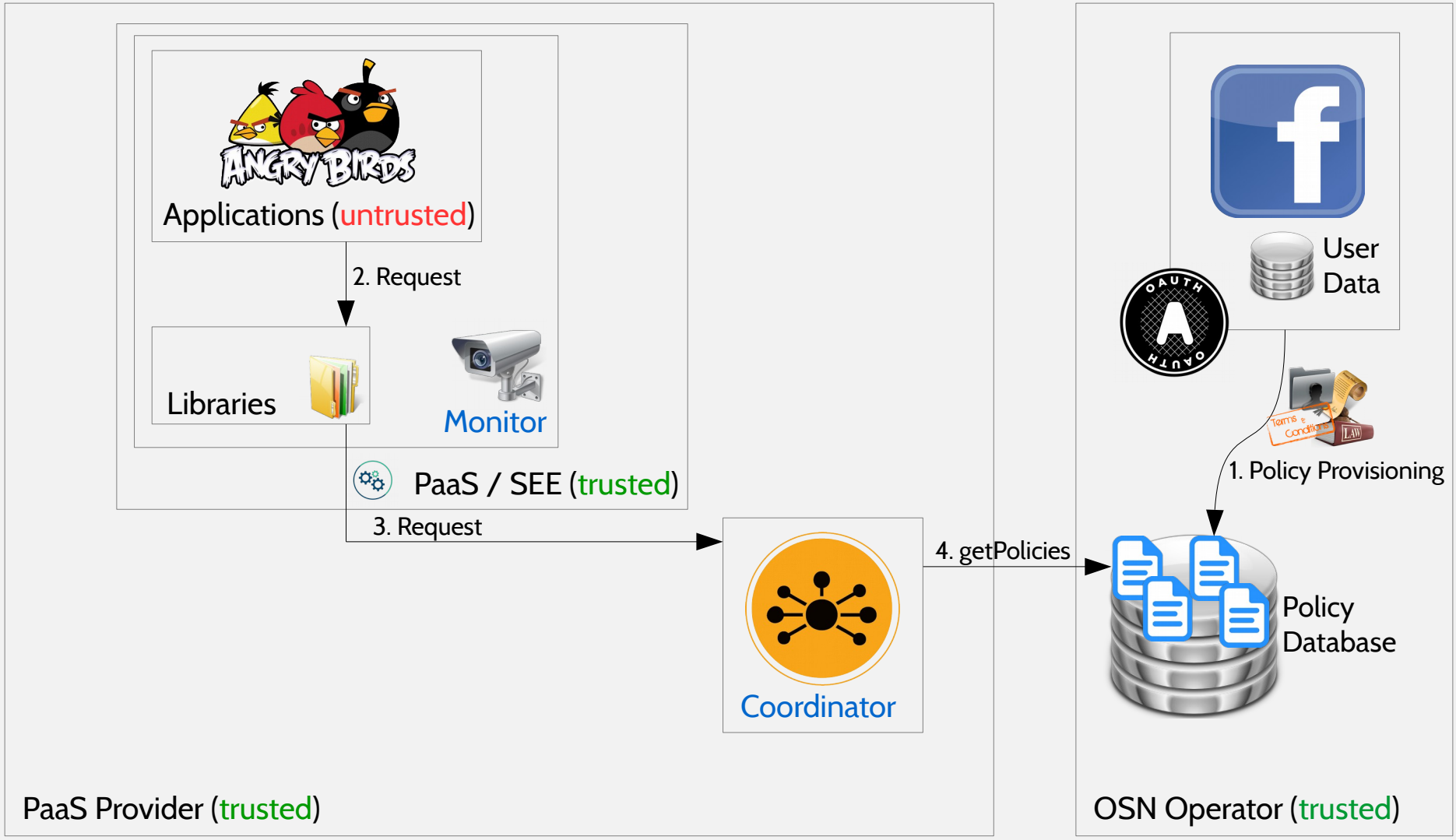
Overview


Applications (untrusted)
2. Request
Libraries
Monitor
PaaS / SEE (trusted)
PaaS Provider (trusted)
User Data
1. Policy Provisioning
Policy Database
OSN Operator (trusted)

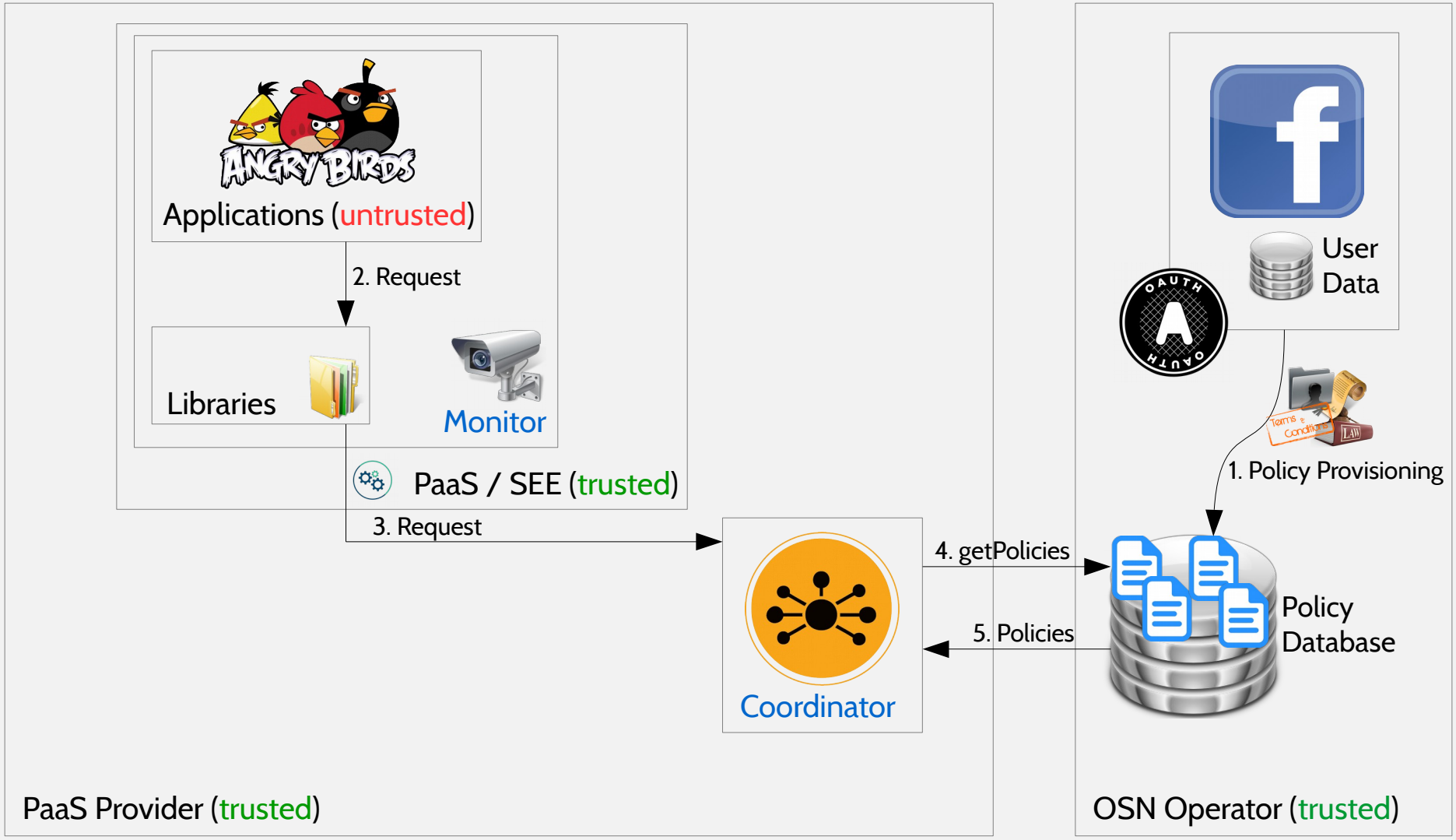
Overview


Applications (untrusted)
2. Request
Libraries
Monitor
PaaS / SEE (trusted)
3. Request
Coordinator
PaaS Provider (trusted)
User Data
1. Policy Provisioning
Policy Database
OSN Operator (trusted)

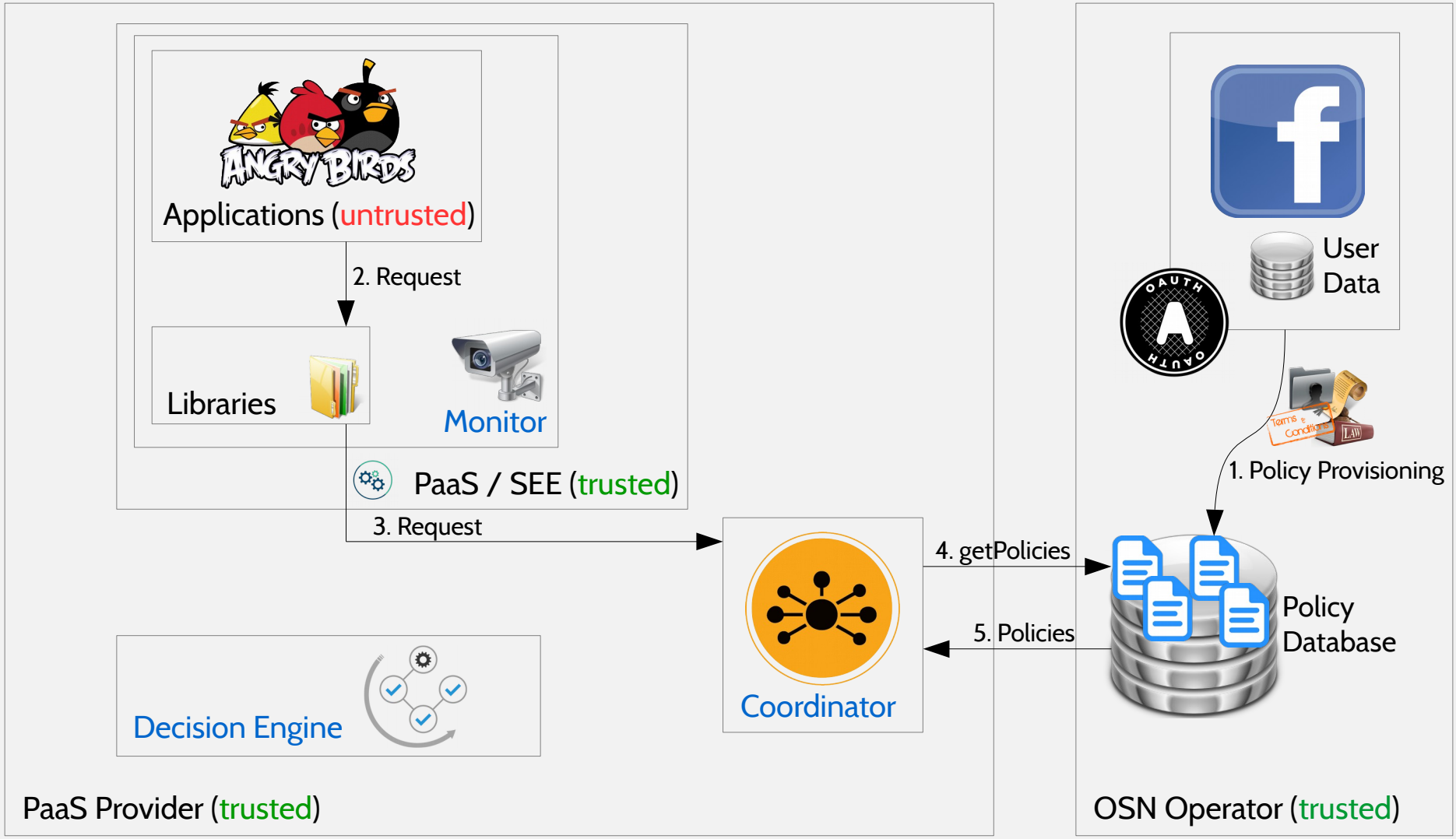
Overview


Applications (untrusted)
2. Request
Libraries
Monitor
PaaS / SEE (trusted)
3. Request
Coordinator
4. getPolicies
PaaS Provider (trusted)
User Data
1. Policy Provisioning
Policy Database
OSN Operator (trusted)

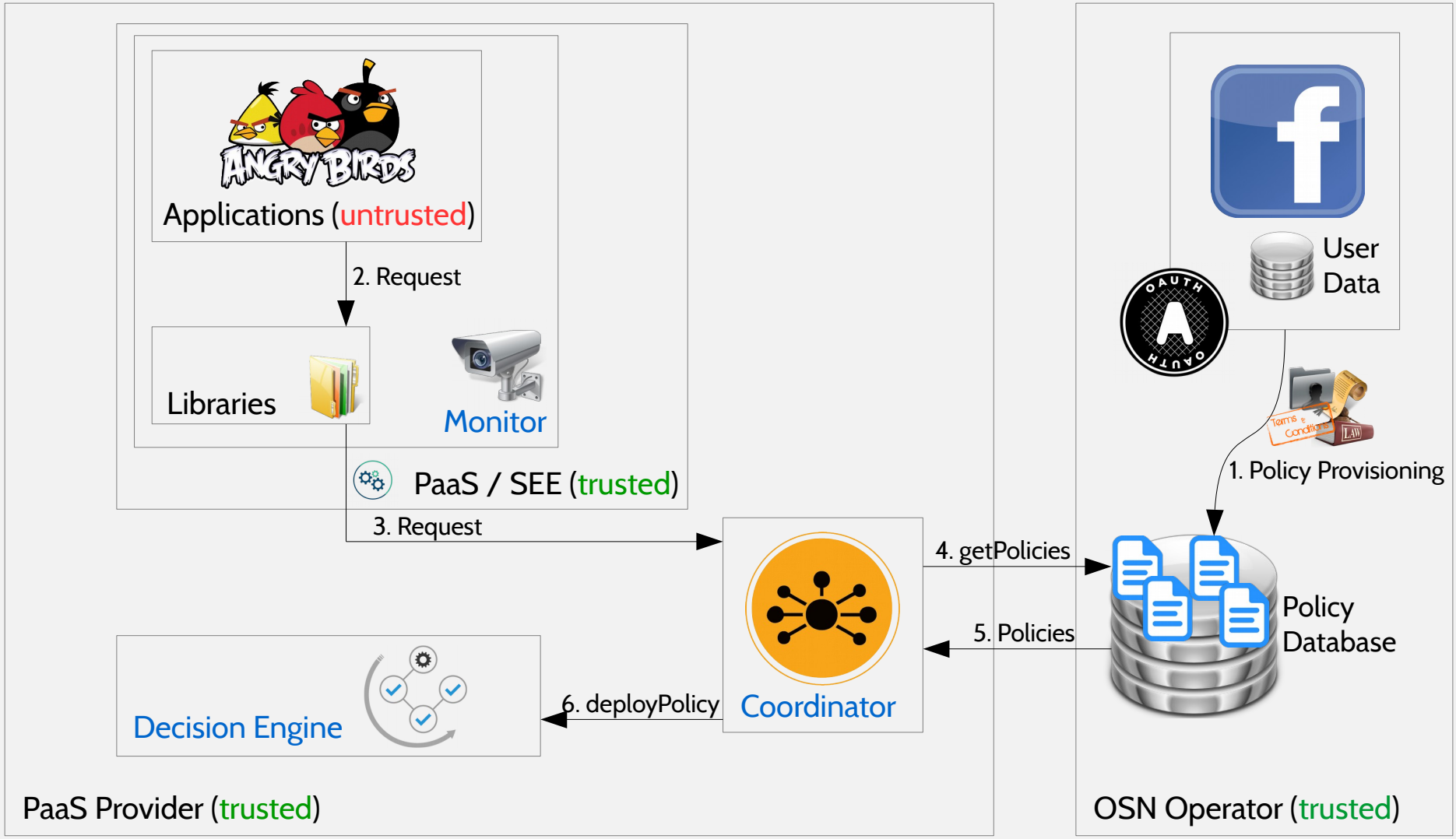
Overview


Applications (untrusted)
2. Request
Libraries
Monitor
PaaS / SEE (trusted)
3. Request
Coordinator
4. getPolicies
5. Policies
PaaS Provider (trusted)
User Data
1. Policy Provisioning
Policy Database
OSN Operator (trusted)

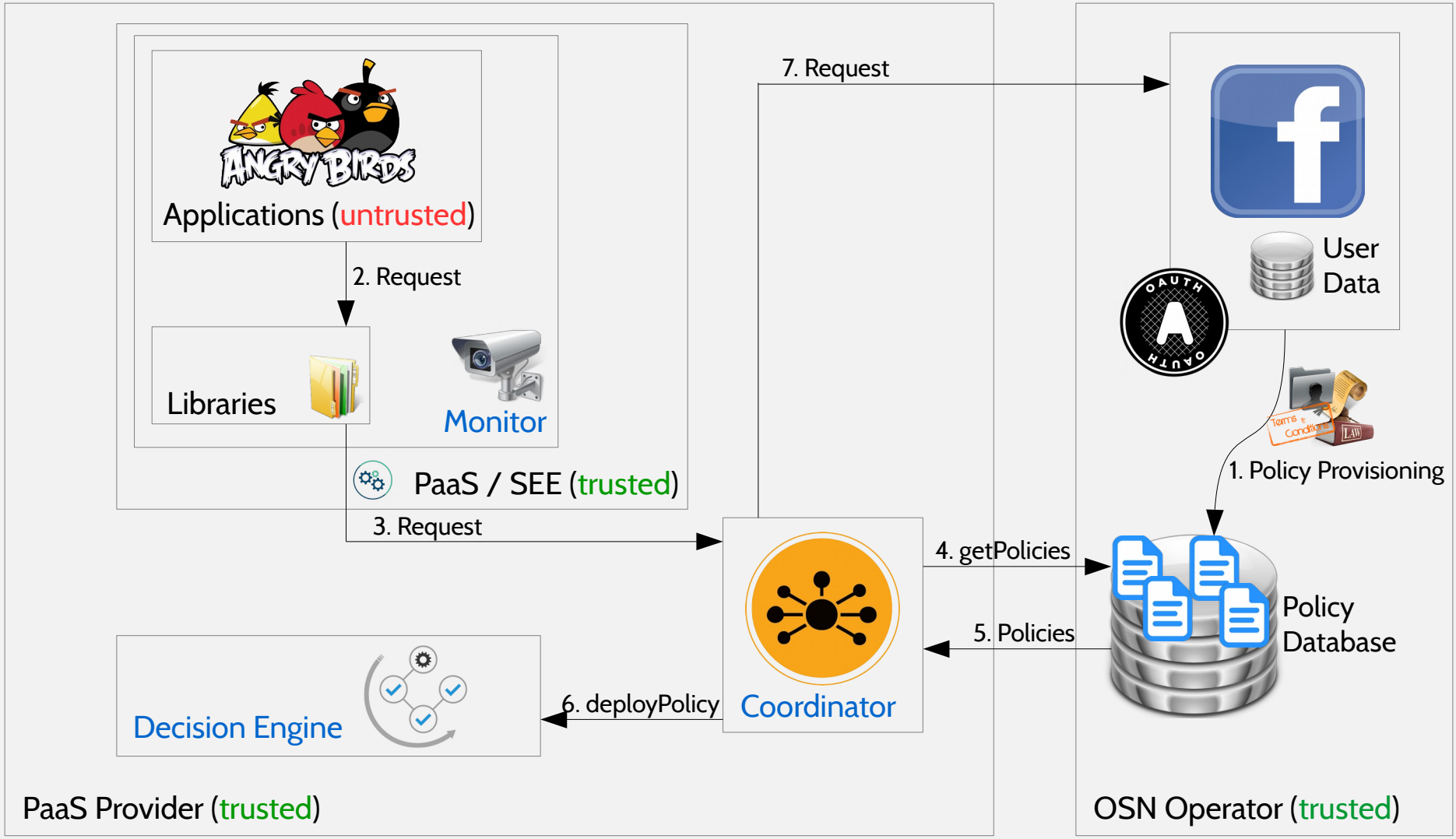
Overview


Applications (untrusted)
2. Request
Libraries
Monitor
PaaS / SEE (trusted)
3. Request
Decision Engine
PaaS Provider (trusted)
Coordinator
4. getPolicies
5. Policies
User Data
1. Policy Provisioning
Policy Database
OSN Operator (trusted)

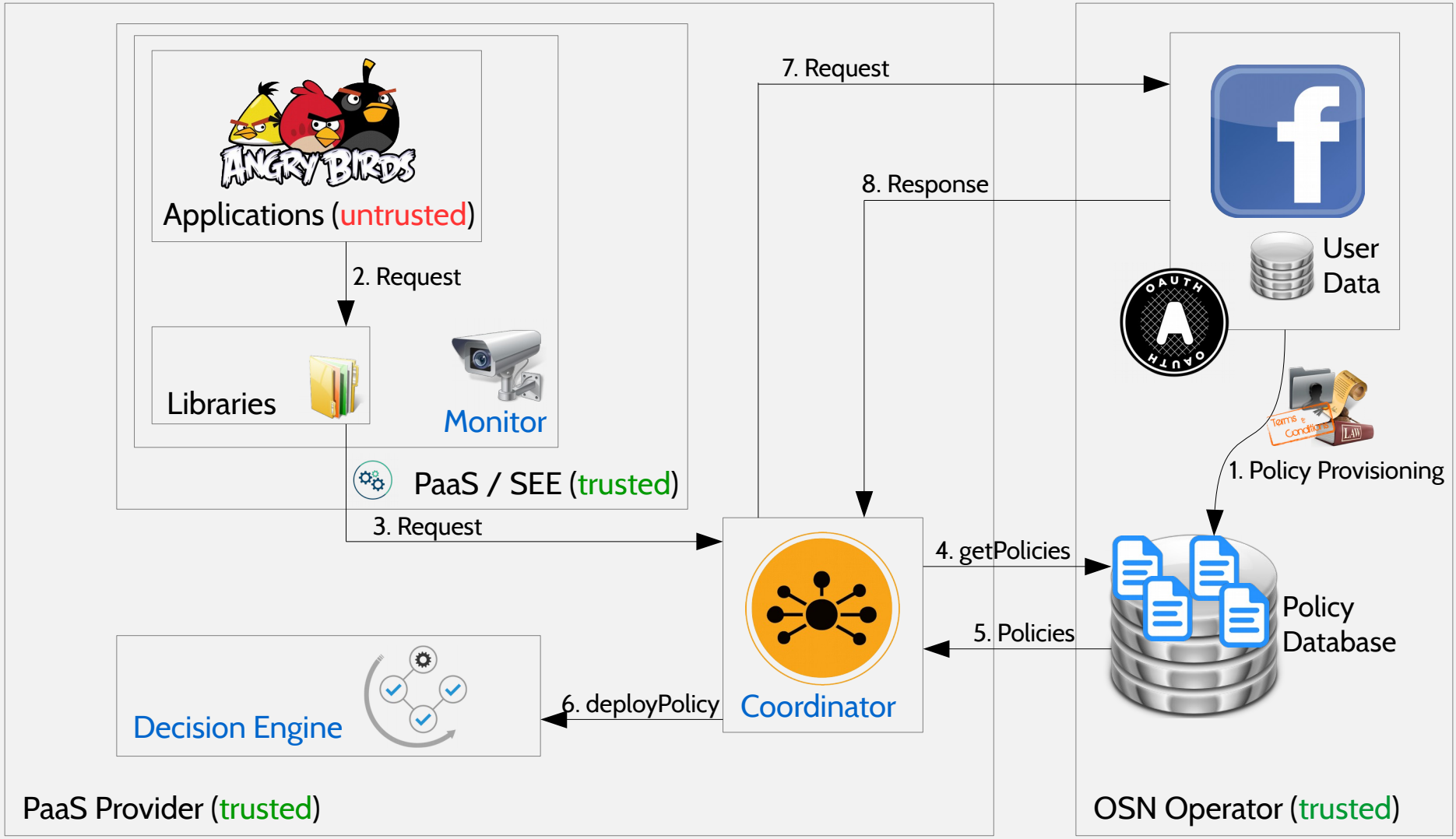
Overview


Applications (untrusted)
2. Request
Libraries
Monitor
PaaS / SEE (trusted)
3. Request
Decision Engine
6. deployPolicy
Coordinator
4. getPolicies
5. Policies
PaaS Provider (trusted)
User Data
1. Policy Provisioning
Policy Database
OSN Operator (trusted)

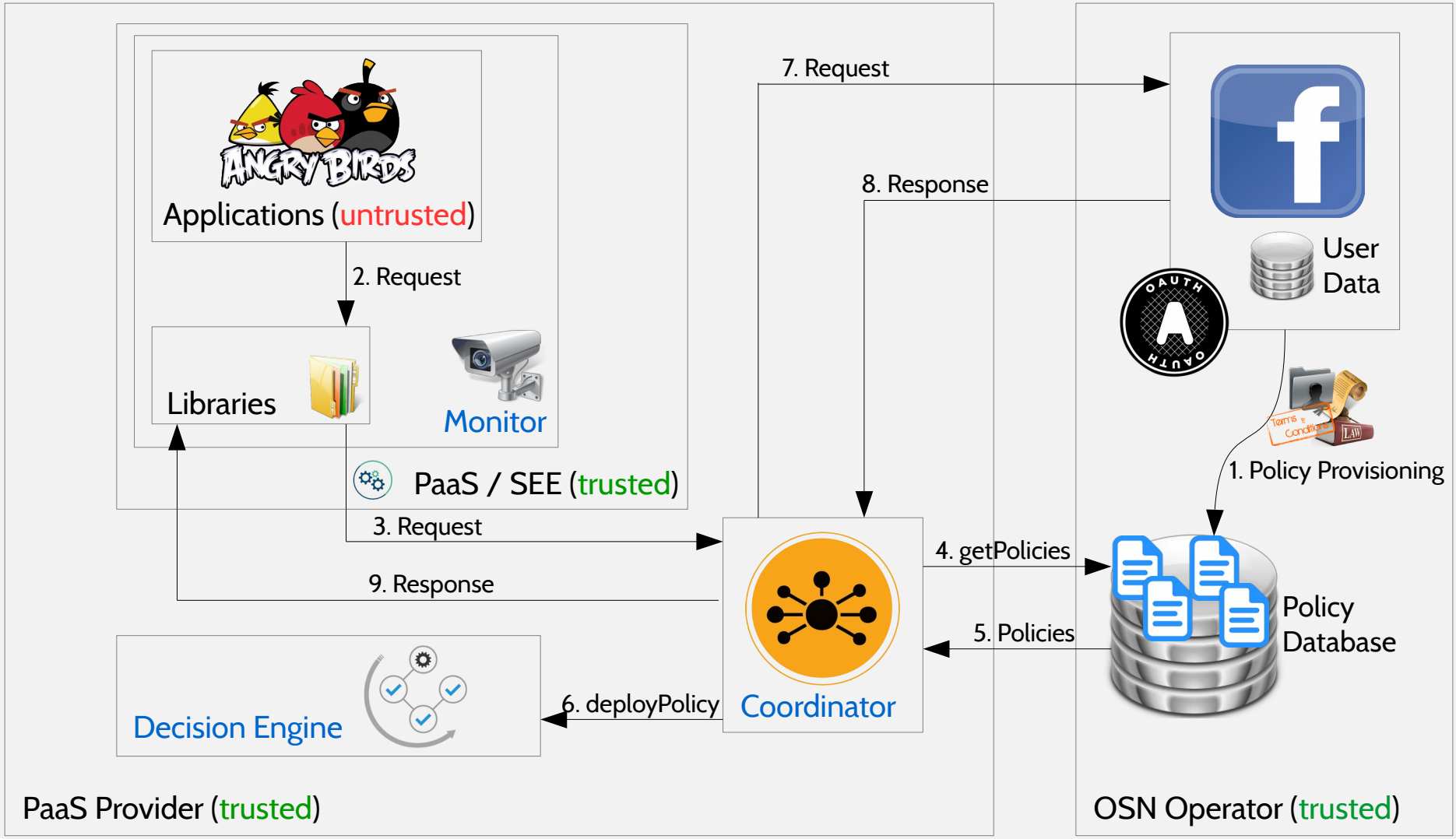
Overview


Applications (untrusted)
2. Request
Libraries
Monitor
PaaS / SEE (trusted)
3. Request
7. Request
User Data
1. Policy Provisioning
4. getPolicies
5. Policies
Policy Database
Coordinator
6. deployPolicy
Decision Engine
PaaS Provider (trusted)
OSN Operator (trusted)

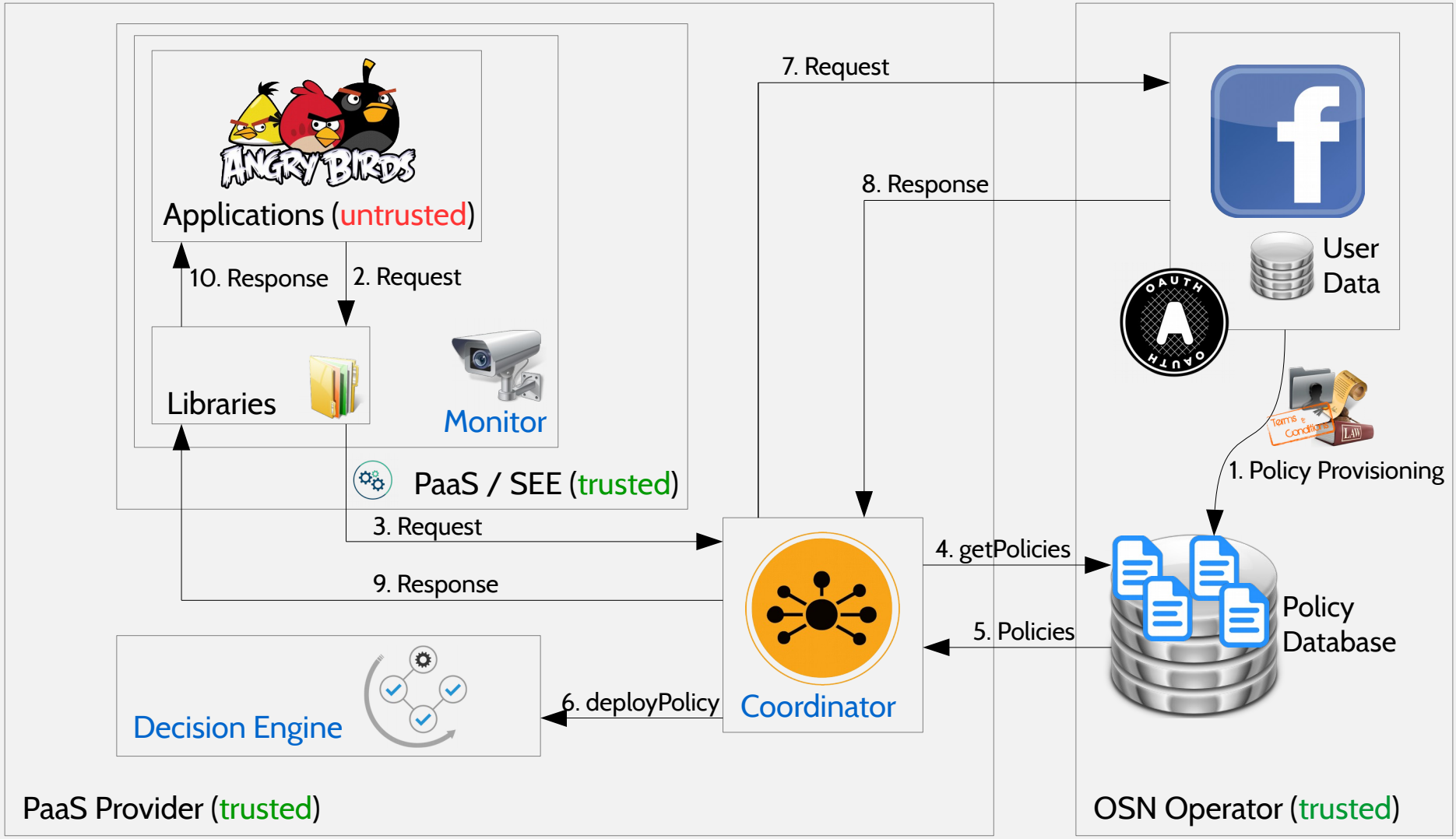
Overview


Applications (untrusted)
2. Request
Libraries
Monitor
PaaS / SEE (trusted)
3. Request
7. Request
8. Response
User Data
1. Policy Provisioning
4. getPolicies
5. Policies
Policy Database
Coordinator
6. deployPolicy
Decision Engine
PaaS Provider (trusted)
OSN Operator (trusted)

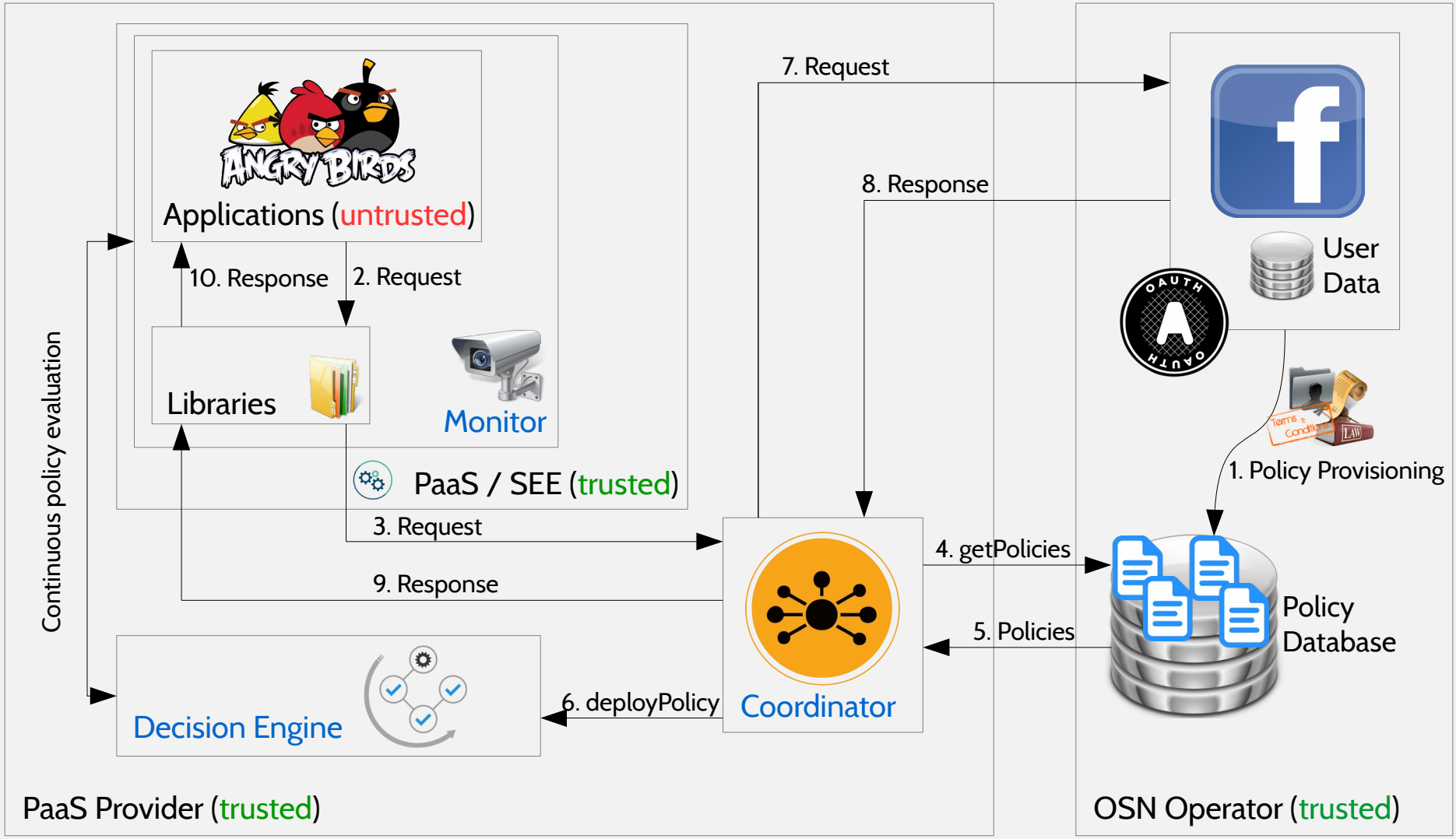
Overview


Applications (untrusted)
2. Request
Libraries
Monitor
PaaS / SEE (trusted)
3. Request
9. Response
Decision Engine
6. deployPolicy
PaaS Provider (trusted)
7. Request
8. Response
Coordinator
4. getPolicies
5. Policies
User Data
1. Policy Provisioning
Policy Database
OSN Operator (trusted)

Overview


Applications (untrusted)
Libraries
Monitor
PaaS / SEE (trusted)
10. Response
2. Request
3. Request
9. Response
7. Request
8. Response
User Data
1. Policy Provisioning
4. getPolicies
5. Policies
Policy Database
Coordinator
6. deployPolicy
Decision Engine
PaaS Provider (trusted)
OSN Operator (trusted)

Overview


Applications (untrusted)
Libraries
Monitor
PaaS / SEE (trusted)
Continuous policy evaluation
10. Response
2. Request
3. Request
9. Response
7. Request
8. Response
1. Policy Provisioning
4. getPolicies
5. Policies
6. deployPolicy
Coordinator
Decision Engine
User Data
Policy Database
PaaS Provider (trusted)
OSN Operator (trusted)

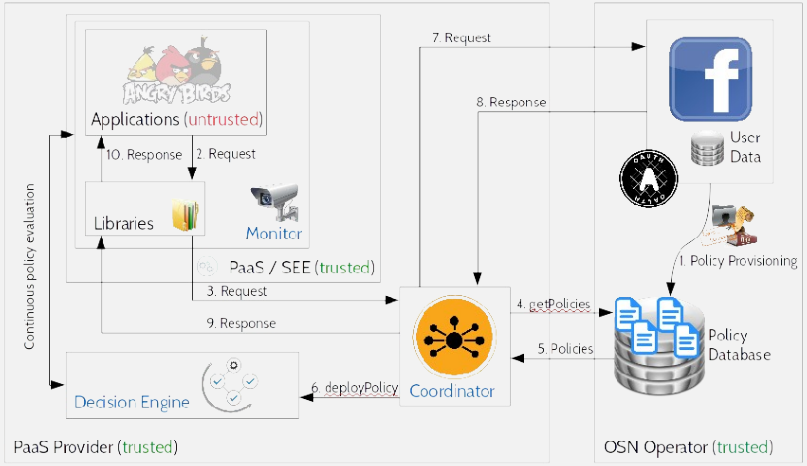
Some details follow …

Some details follow …

- Policy Provisioning

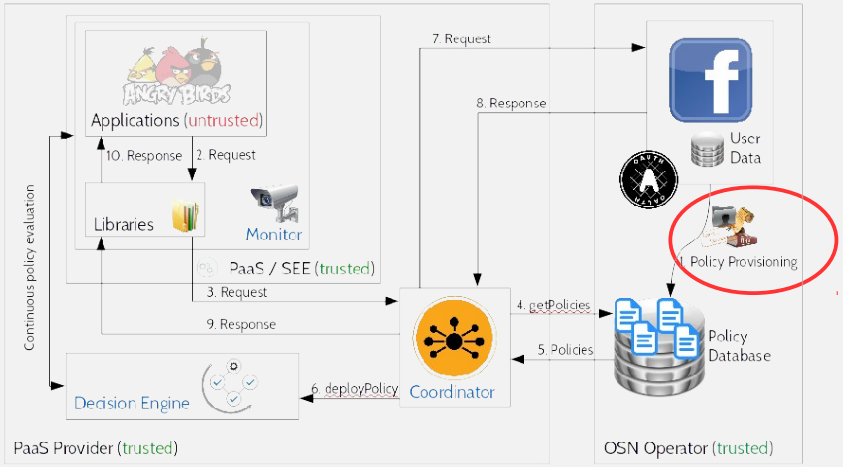
Some details follow …

Policy Provisioning



Some details follow …

Policy Provisioning



Policy Provisioning

Policy Provisioning

“If you cache data you receive from us, [. . .] keep it up to date”

Policy Provisioning

“If you cache data you receive from us, [. . .] keep it up to date”

→

Policy Provisioning

“If you cache data you receive from us, [. . .] keep it up to date”

→

“Whenever some data is processed by the application, then it must have been received from the OSN within the last 24 hours”

Policy Provisioning

"If you cache data you receive from us, [. . .] keep it up to date"

→

"Whenever some data is processed by the application, then it must have been received from the OSN within the last 24 hours"

→

Policy Provisioning

"If you cache data you receive from us, [. . .] keep it up to date"

→

"Whenever some data is processed by the application, then it must have been received from the OSN within the last 24 hours"

→

Event:

Condition:

Action:

Policy Provisioning

"If you cache data you receive from us, [. . .] keep it up to date"

→

"Whenever some data is processed by the application, then it must have been received from the OSN within the last 24 hours"

→

Event: process(data)

Condition:

Action:

Policy Provisioning

"If you cache data you receive from us, [. . .] keep it up to date"

→

"Whenever some data is processed by the application, then it must have been received from the OSN within the last 24 hours"

→

Event: process(data)

Condition:

Action: <inhibit>

Policy Provisioning

"If you cache data you receive from us, [. . .] keep it up to date"

→

"Whenever some data is processed by the application, then it must have been received from the OSN within the last 24 hours"

→

Event: process(data)

Condition: not(repmin(24[hours], 1, receive(data)))

Action: <inhibit>

Policy Provisioning

"If you cache data you receive from us, [. . .] keep it up to date"

→

"Whenever some data is processed by the application, then it must have been received from the OSN within the last 24 hours"

→

Event: process(data)

Condition: not(repmin(24[hours], 1, receive(data)))

Action: <inhibit>

Complex LTL formulas:
- propositional
- temporal
- cardinal
- spatial

constraints

Policy Provisioning

"If you cache data you receive from us, [...] keep it up to date"

$$
\begin{aligned}
\Psi &= \underline{false} \mid \mathcal{E} \\
\Phi^{\Sigma} &= \underline{isNotIn}(\mathcal{D}, \mathbb{P}(\mathcal{C})) \mid \underline{isCombined}(\mathcal{D}, \mathcal{D}, \mathbb{P}(\mathcal{C})) \mid \underline{isMaxIn}(\mathcal{D}, \mathbb{N}, \mathbb{P}(\mathcal{C})) \\
\Phi &= (\Phi) \mid \Psi \mid \Phi^{\Sigma} \mid \Phi \ \underline{and}\ \Phi \mid \underline{not}(\Phi) \mid \Phi \ \underline{since}\ \Phi \mid \Phi \ \underline{before}\ \mathbb{N} \mid \underline{repmin}(\mathbb{N}, \mathbb{N}, \mathcal{E})
\end{aligned}
$$

$$
\begin{aligned}
&\forall t \in Trace, i \in \mathbb{N}, \varphi \in \Phi \bullet (t,i) \models \varphi \iff (\varphi \neq \underline{false}) \wedge \\
&\quad \exists e \in \mathcal{E}, e' \in t(i) \bullet (\varphi = e \wedge (e', \sigma_t^i)\ refines_{\Sigma}\ e) \\
&\vee \exists d \in \mathcal{D}, C \subseteq \mathcal{C} \bullet (\varphi = \underline{isNotIn}(d, C) \wedge \forall c \in C \bullet d \notin \sigma_t^i(c)) \\
&\vee \exists d_1, d_2 \in \mathcal{D}, C \subseteq \mathcal{C} \bullet (\varphi = \underline{isCombined}(d_1, d_2, C) \wedge \exists c \in C \bullet \{d_1, d_2\} \subseteq \sigma_t^i(c)) \\
&\vee \exists d \in \mathcal{D}, m \in \mathbb{N}, C \subseteq C \bullet (\varphi = \underline{isMaxIn}(d, m, C) \wedge |\{c \in C | d \in \sigma_t^i(c)\}| \leq m) \\
&\vee \exists \alpha, \beta \in \Phi \bullet ((\varphi = \underline{not}(\alpha) \wedge \neg((t,i) \models \alpha)) \\
&\qquad \vee (\varphi = \alpha\ \underline{and}\ \beta \wedge (t,i) \models \alpha \wedge (t,i) \models \beta) \\
&\qquad \vee (\varphi = \alpha\ \underline{or}\ \beta \wedge (t,i) \models \alpha \vee (t,i) \models \beta) \\
&\qquad \vee (\varphi = \alpha\ \underline{since}\ \beta \wedge \exists j \in [0,i] \bullet ((t,j) \models \beta \wedge \forall k \in (j,i] \bullet (t,k) \models \alpha) \\
&\qquad\quad \vee \forall k \in [0,i] \bullet (t,k) \models \alpha)) \\
&\vee \exists \alpha \in \Phi, j \in \mathbb{N} \bullet (\varphi = \alpha\ \underline{before}\ j \wedge (t, i-j) \models \alpha) \\
&\vee \exists j, m \in \mathbb{N}, e \in \mathcal{E} \bullet (\varphi = \underline{repmin}(j, m, e) \\
&\qquad \wedge m \leq \textstyle\sum_{k=0}^{j-1} |\{e' \in t(i-k) | (e', \sigma_t^{i-k})\ refines_{\Sigma}\ e\}|)
\end{aligned}
$$

ex LTL formulas:
positional
poral
inal
ial
constraints

Event:

Condition: not(repmin(24[hours], 1, receive(data)))

Action: <inhibit>

Some details follow …

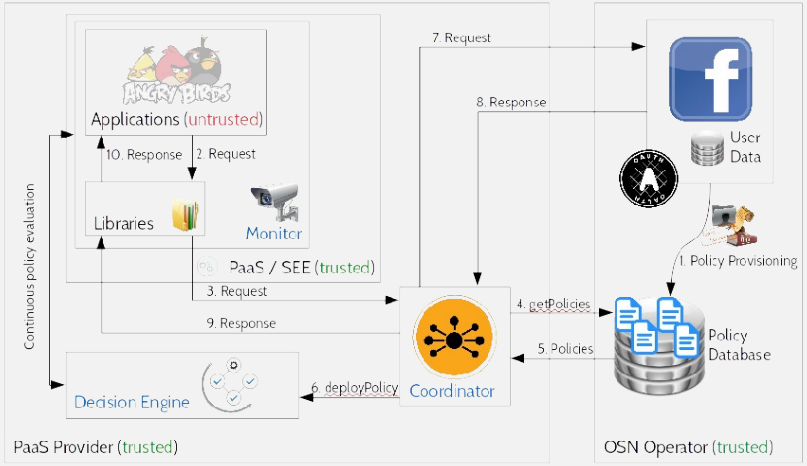
Policy Provisioning

Application Deployment

Some details follow …

Policy Provisioning

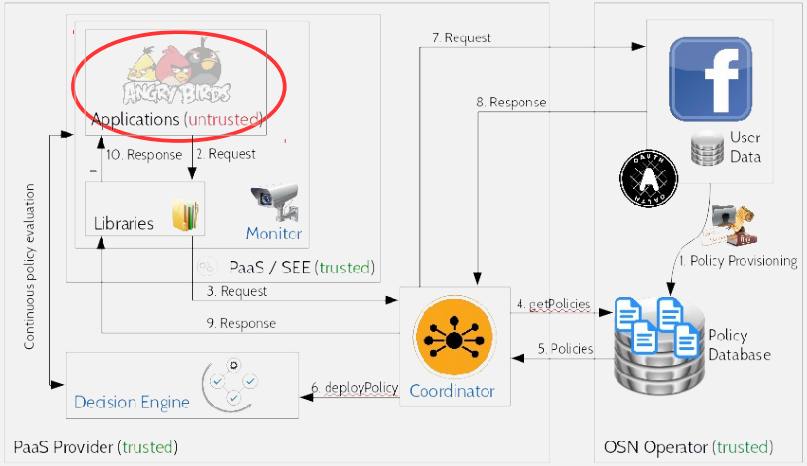
Application Deployment



Some details follow …

Policy Provisioning

Application Deployment



Application Deployment

Application Deployment

Analysis of binary app to find

Application Deployment

Analysis of binary app to find

Data sources

e.g., retrieve from OSN

```
Sources:
  Source36:
    Location: Common.updateCurrentUser(...)V:61
    Signature: com.restfb.DefaultFacebookClient.
               fetchObject(Ljava/lang/String;
               Ljava/lang/Class;[Lcom/restfb/Parameter;
          )Ljava/lang/Object;
    Return
```

Application Deployment

Analysis of binary app to find

- Data sources
  - e.g., retrieve from OSN
- Data sinks
  - e.g., data usage/sharing

```
Sources:
  Source36:
    Location: Common.updateCurrentUser(...)V:61
    Signature: com.restfb.DefaultFacebookClient.
               fetchObject(Ljava/lang/String;
               Ljava/lang/Class;[Lcom/restfb/Parameter;
          )Ljava/lang/Object;
    Return

Sinks:
  Sink15:
    Location: Main.loggedInWork(...)V:153
    Signature: org.apache.catalina.connector.
               CoyoteWriter.println(Ljava/lang/String;)V
    ParamIndex: 1
```

Application Deployment

Analysis of binary app to find

- Data sources
  - e.g., retrieve from OSN
- Data sinks
  - e.g., data usage/sharing
- Dependencies between them

```
Sources:
  Source36:
    Location: Common.updateCurrentUser(...)V:61
    Signature: com.restfb.DefaultFacebookClient.
               fetchObject(Ljava/lang/String;
               Ljava/lang/Class;[Lcom/restfb/Parameter;
          )Ljava/lang/Object;
    Return

Sinks:
  Sink15:
    Location: Main.loggedInWork(...)V:153
    Signature: org.apache.catalina.connector.
               CoyoteWriter.println(Ljava/lang/String;)V
    ParamIndex: 1

Flows:
  Sink15 --> Source36
```

Application Deployment

Analysis of binary app to find

- Data sources
  - e.g., retrieve from OSN
- Data sinks
  - e.g., data usage/sharing
- Dependencies between them

```
Sources:
  Source36:
    Location: Common.updateCurrentUser(...)V:61
    Signature: com.restfb.DefaultFacebookClient.
               fetchObject(Ljava/lang/String;
               Ljava/lang/Class;[Lcom/restfb/Parameter;
          )Ljava/lang/Object;
    Return

Sinks:
  Sink15:
    Location: Main.loggedInWork(...)V:153
    Signature: org.apache.catalina.connector.
               CoyoteWriter.println(Ljava/lang/String;)V
    ParamIndex: 1

Flows:
  Sink15 --> Source36
```

Instrumentation of sources and sinks for

Application Deployment

Analysis of binary app to find

- Data sources
  - e.g., retrieve from OSN
- Data sinks
  - e.g., data usage/sharing
- Dependencies between them

```
Sources:
  Source36:
    Location: Common.updateCurrentUser(...)V:61
    Signature: com.restfb.DefaultFacebookClient.
               fetchObject(Ljava/lang/String;
               Ljava/lang/Class;[Lcom/restfb/Parameter;
          )Ljava/lang/Object;
    Return

Sinks:
  Sink15:
    Location: Main.loggedInWork(...)V:153
    Signature: org.apache.catalina.connector.
               CoyoteWriter.println(Ljava/lang/String;)V
    ParamIndex: 1

Flows:
  Sink15 --> Source36
```

Instrumentation of sources and sinks for

- Data flow tracking
- Policy decisions

Some details follow …

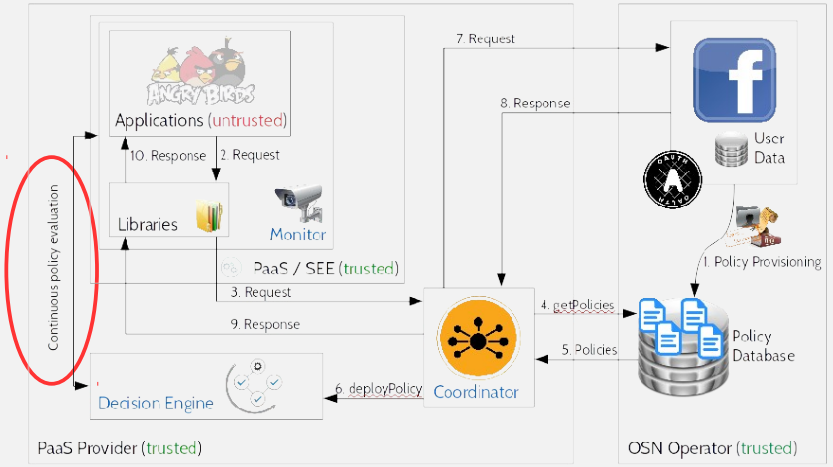
Policy Provisioning

Application Deployment

Event signaling for

data flow tracking

policy decisions



Event Signaling

Event Signaling

Signal invocations of

- data sources
- data sinks

Event Signaling

Signal invocations of

- data sources
- data sinks

to decision engine:

Event Signaling

Signal invocations of

- data sources
- data sinks

to decision engine:

- Is data read from source?
- Is data written to sink?

Event Signaling

Signal invocations of

- data sources
- data sinks

to decision engine:

- Is data read from source?
- Is data written to sink?

} i.e., does event of ECA rule match?

Event:	process(data)
Condition:	not(repmin(24[hours], 1, receive(data)))
Action:	<inhibit>

Event Signaling

Signal invocations of

data sources

data sinks

to decision engine:

Is data read from source?
Is data written to sink?
} i.e., does event of ECA rule match?

If Yes: Evaluate condition

Event:	process(data)
Condition:	not(repmin(24[hours], 1, receive(data)))
Action:	<inhibit>

Event Signaling

Signal invocations of

- data sources
- data sinks

to decision engine:

- Is data read from source?
- Is data written to sink?

} i.e., does event of ECA rule match?

- If Yes: Evaluate condition
  - If True: Apply action

Event:	process(data)
Condition:	not(repmin(24[hours], 1, receive(data)))
Action:	<inhibit>

Event Signaling

Signal invocations of

data sources

data sinks

to decision engine:

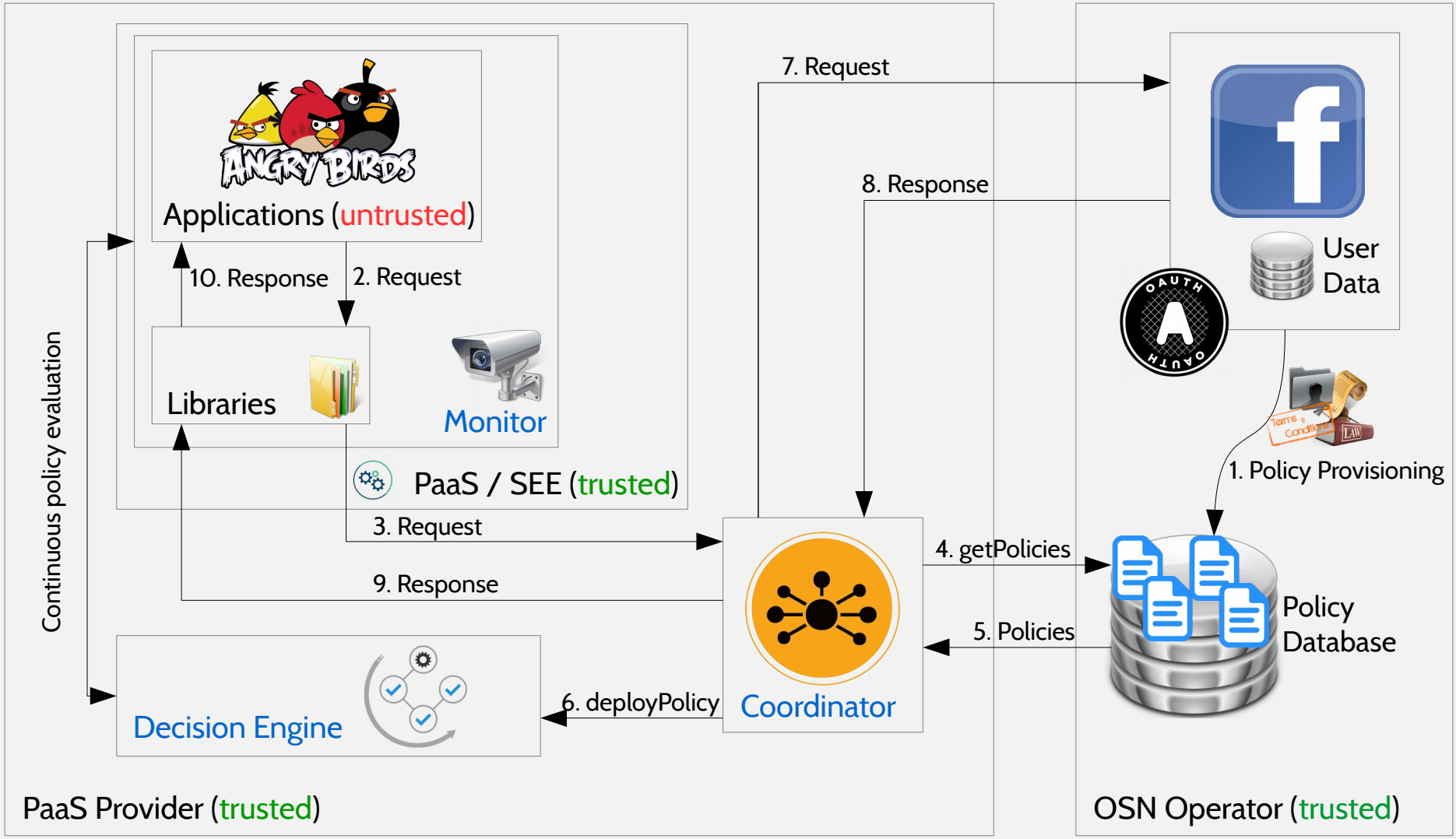
Is data read from source?
Is data written to sink?
} i.e., does event of ECA rule match?

If Yes: Evaluate condition

If True: Apply action

Event:	process(data)
Condition:	not(repmin(24[hours], 1, receive(data)))
Action:	<inhibit>

Overview


Applications (untrusted)
Libraries
Monitor
PaaS / SEE (trusted)
Continuous policy evaluation
10. Response
2. Request
3. Request
9. Response
7. Request
8. Response
1. Policy Provisioning
4. getPolicies
5. Policies
6. deployPolicy
Coordinator
Decision Engine
User Data
Policy Database
PaaS Provider (trusted)
OSN Operator (trusted)

Evaluation

Evaluation

Between 15% and 41% performance overhead

Evaluation

Between 15% and 41% performance overhead

Depends much on the application and policy

Evaluation

Between 15% and 41% performance overhead

Depends much on the application and policy

Problem: Real-world apps are not available

Summary

Summary

Protection from data misuse is possible

Summary

Protection from data misuse is possible

Critical requirements

Summary

- Protection from data misuse is possible
- Critical requirements
  - User awareness

Summary

Protection from data misuse is possible

Critical requirements

- User awareness
- Transparency for all involved parties

Summary

Protection from data misuse is possible

Critical requirements

- User awareness
- Transparency for all involved parties