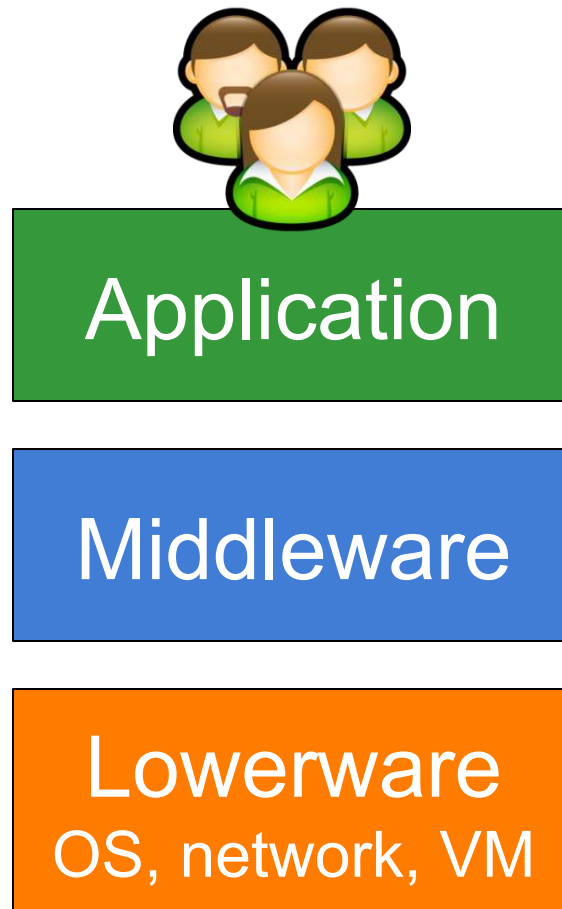


What is Middleware Made Of?

François Taïani, Jackie Rice, Paul Rayson

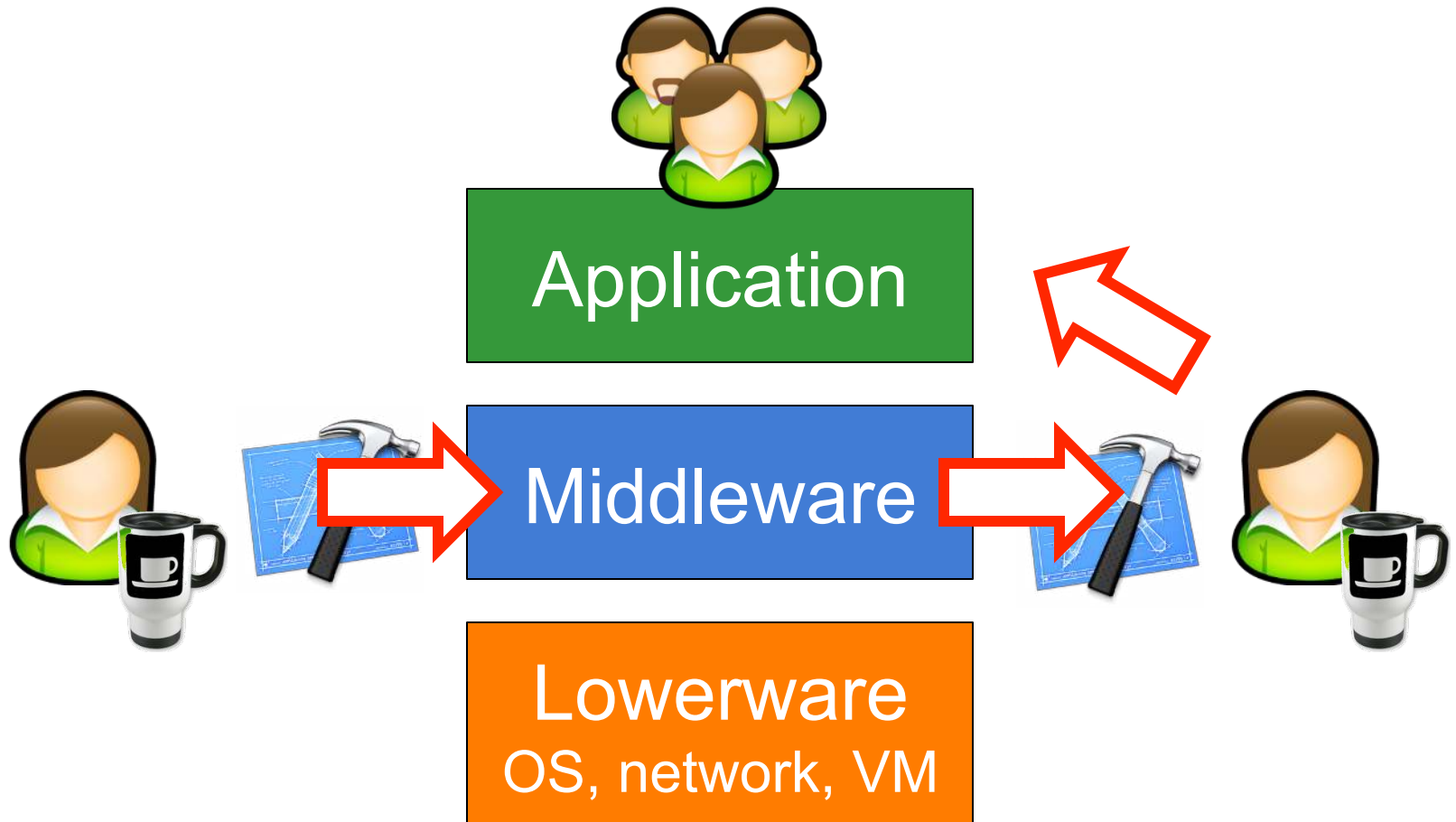


What's Middleware for?



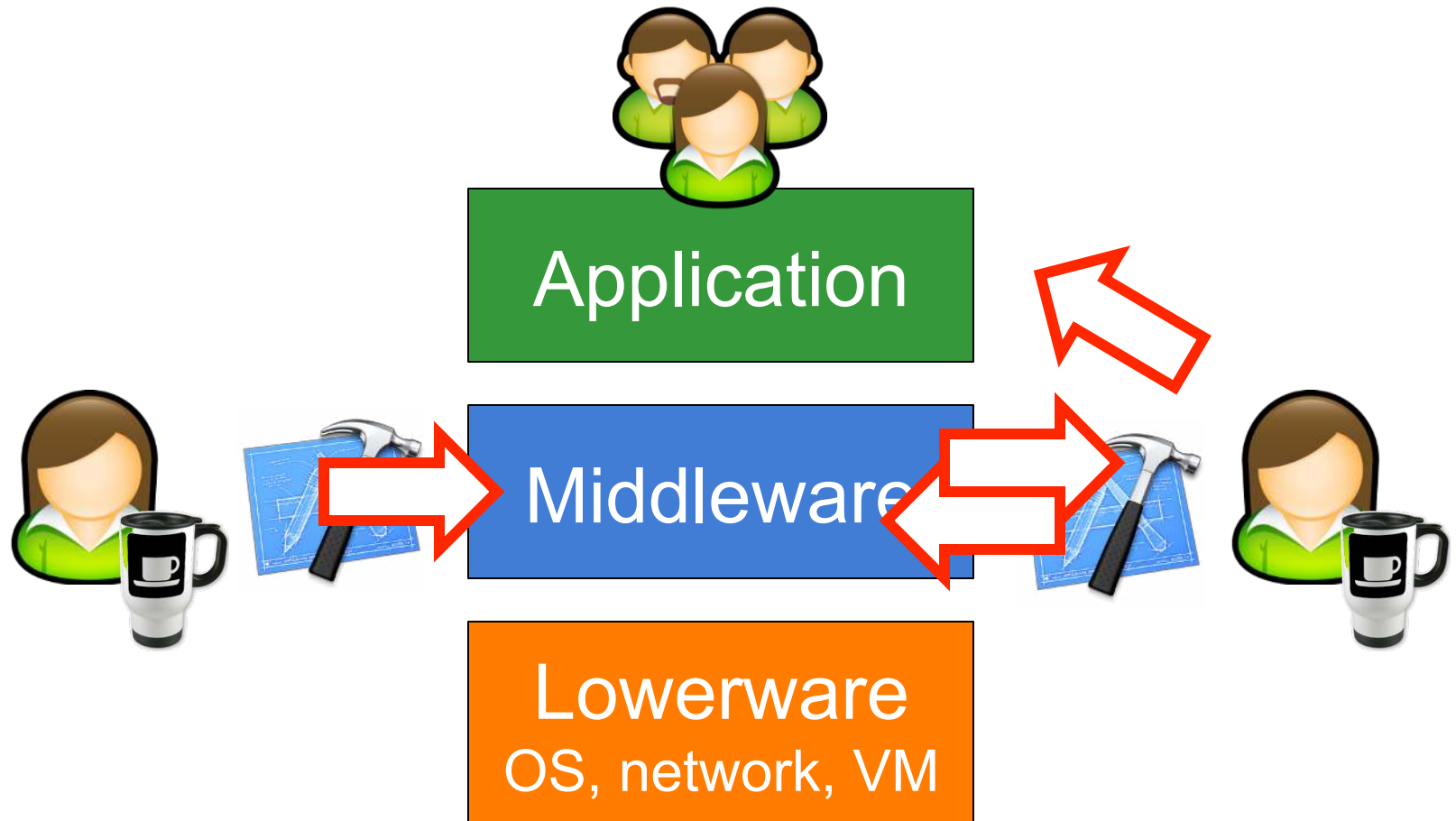
What's Middleware for?

- Both a vertical and horizontal link



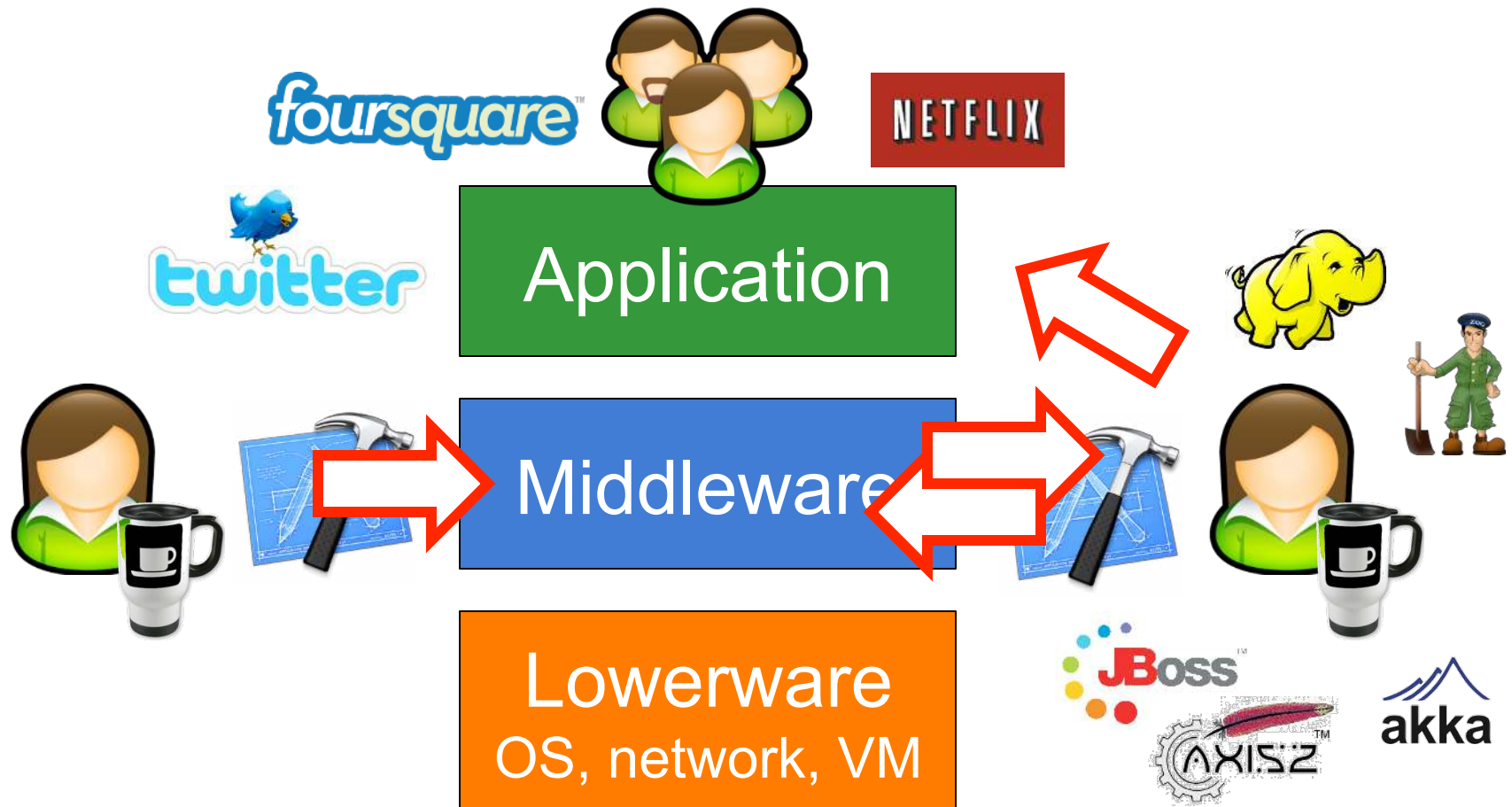
What's Middleware for?

- The link goes both ways for open source software

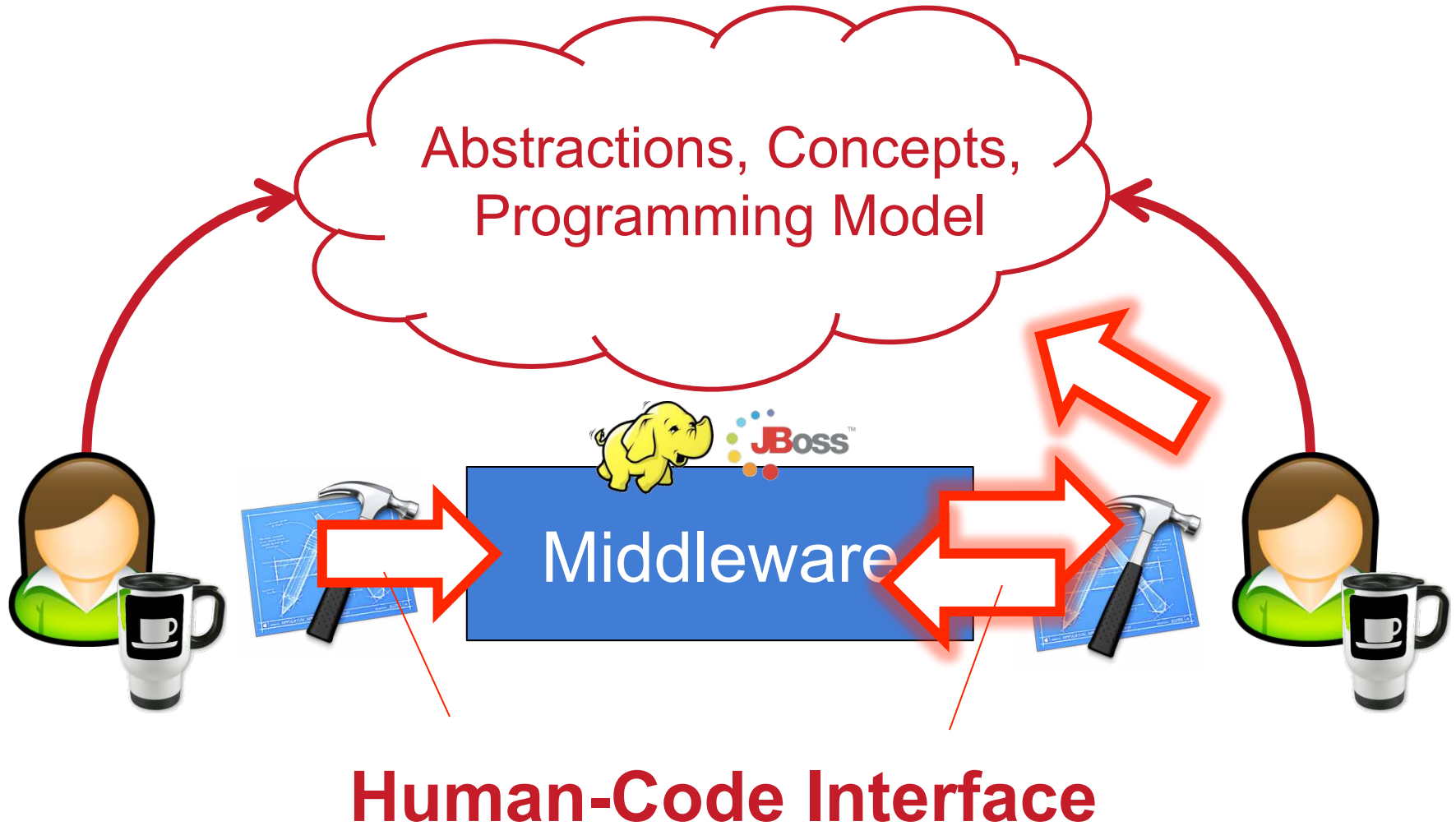


What's Middleware for?

- The link goes both ways for open source software



Middleware as H2H Dialogue



Our Problem

What is the **state of abstractions**
in production middleware?

Our Problem

What is the **state of abstractions**
in production middleware?

impact?

research?

training?

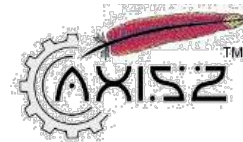
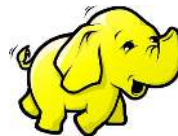
adoption?

Intuition

- Use **program identifiers** to extract concepts
 - convey meaning and intent
 - reflect key concepts and abstractions

CMPCClusteredInMemoryPersistenceManager

- Goal: **Comparative Study** across 4 MW products
 - common conceptual foundation?
 - emerging trends?
 - differences?



Identifiers Convey Meaning

```
public Object dg7(r8s a4t) { [..]
    Object[] ruy = {a4t};
    String[] ewo = {"aj3.c5y.a4t.r8s"};
    Object wu7 = fp1.dg7(af3, "dg7", ruy, ewo);

    Long uds = (Long) a4t.getValue("AHG");
    fy8 as2 = new fy8();
    if (uds.longValue() != ds0.ds2()) {
        as2.l5e = new ArrayList<Object>(ds0.z9p());
        as2.h2x = ds0.ds2();
    }
    as2.x1t = wu7; [..]
    return as2;
}
```

???



HA Replication

Identifiers Convey Meaning

```
public Object invoke(Invocation invocation) { [...]
    Object[] args = {invocation};
    String[] sig = {"org.jboss.invocation.Invocation"};
    Object rtn = mbeanServer.invoke(targetName, "invoke",
    [...]

    Long clientId = (Long) invocation.getValue("CLIENT_VIEW_ID");
    HARMIResponse rsp = new HARMIResponse();
    if (clientId.longValue() != target.getCurrentViewId()) {
        rsp.newReplicants = new ArrayList<Object>(target.getReplicantList());
        rsp.currentViewId = target.getCurrentViewId();
    }
    rsp.response = rtn; [...]
    return rsp;
}
```

Got it ✓

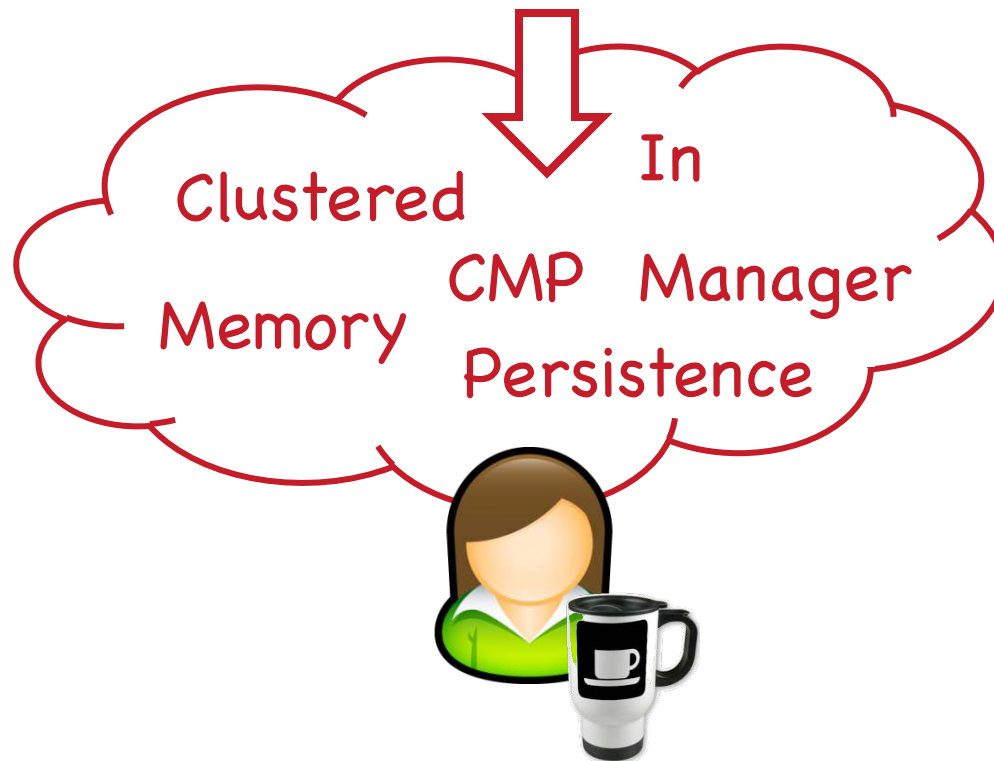


HA Replication

Approach: Software Linguistics

- Growing set of techniques in experimental SE
 - Here: decompose identifiers in their word parts

CMPClusteredInMemoryPersistenceManager



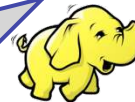
Metrics

CMP Clustered In Memory Persistence Manager

10 locs

Task Memory Manager Thread

5 locs



■ For Concepts

→ **Class Frequency:** Memory: 2, CMP 1

→ **LoCs Weight:** Manager: 15, Persistence: 10, Task 5

→ **Sharing:** Manager 2 projects, Task 1

■ For projects

→ **Name Length:** JBoss 6

→ **Conceptual Footprint:** Hadoop 4

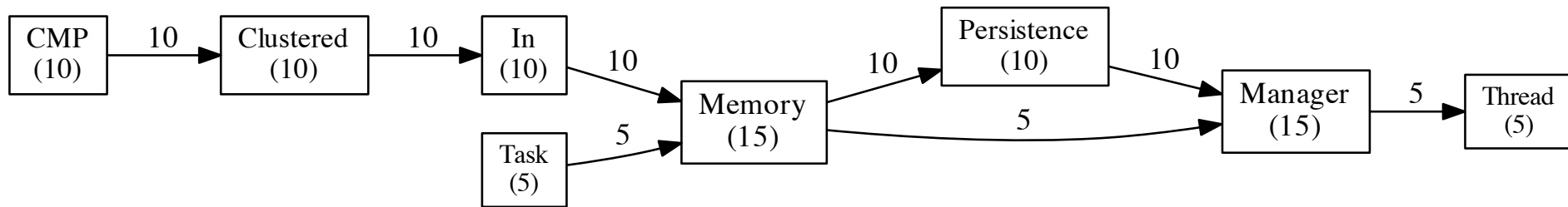
→ **Sharing:** JBoss 2 6, Hadoop 2 4

Collocation Analysis

■ Concept Graph

CMP → Clustered → In → Memory → Persistence → Manager

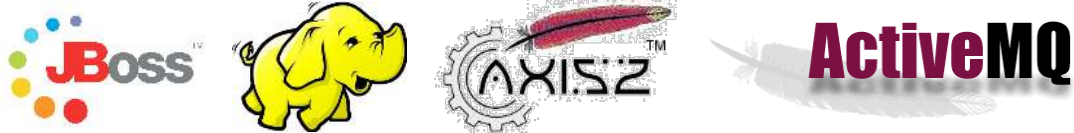
Task → Memory → Manager → Thread



■ (Counting Repeating Patterns (“n-grams”))

→ e.g. “Meta Data”, “Input File”)

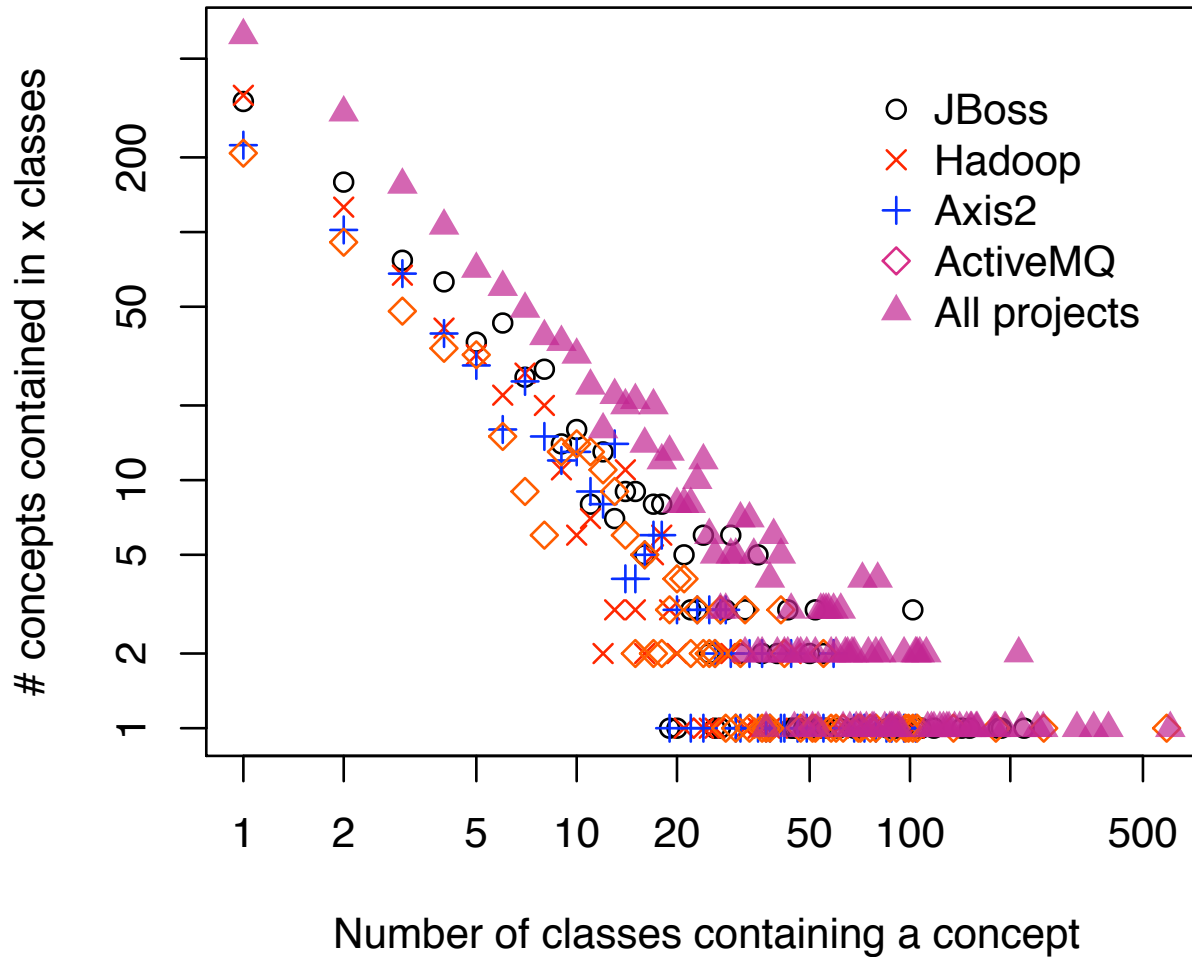
Targets



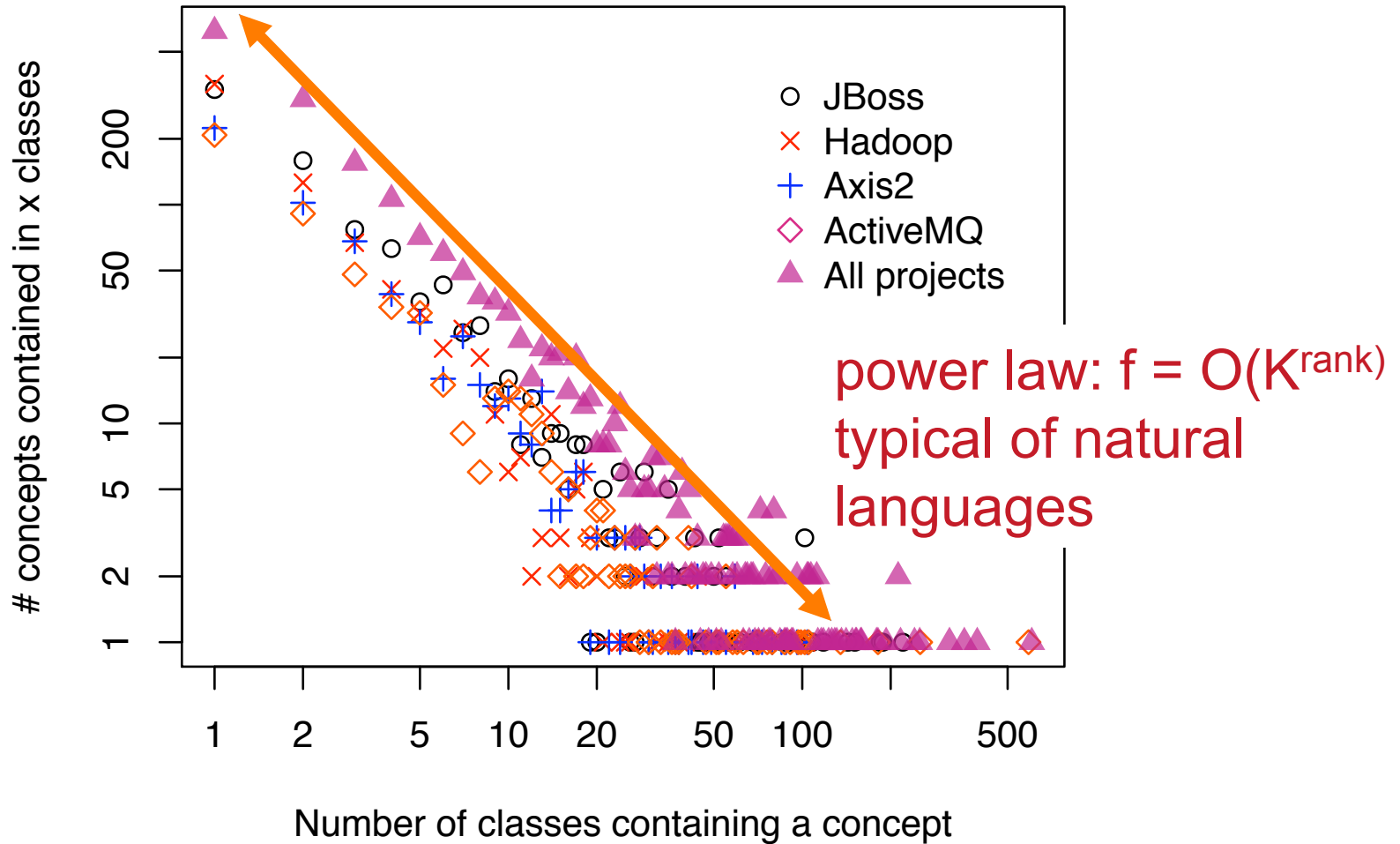
<i>middleware</i>	<i>LoCs</i>	<i># classes</i>
JBoss AS (6.0.0)	345,063	2,290
Hadoop (1.0.3)	255,563	1,327
Axis2 (1.6.2)	302,315	1,514
ActiveMQ (5.6.0)	247,801	1,873
Total	1,150,742	7,004

- Mature, substantial POJO projects
- Different services, different projects, different archi

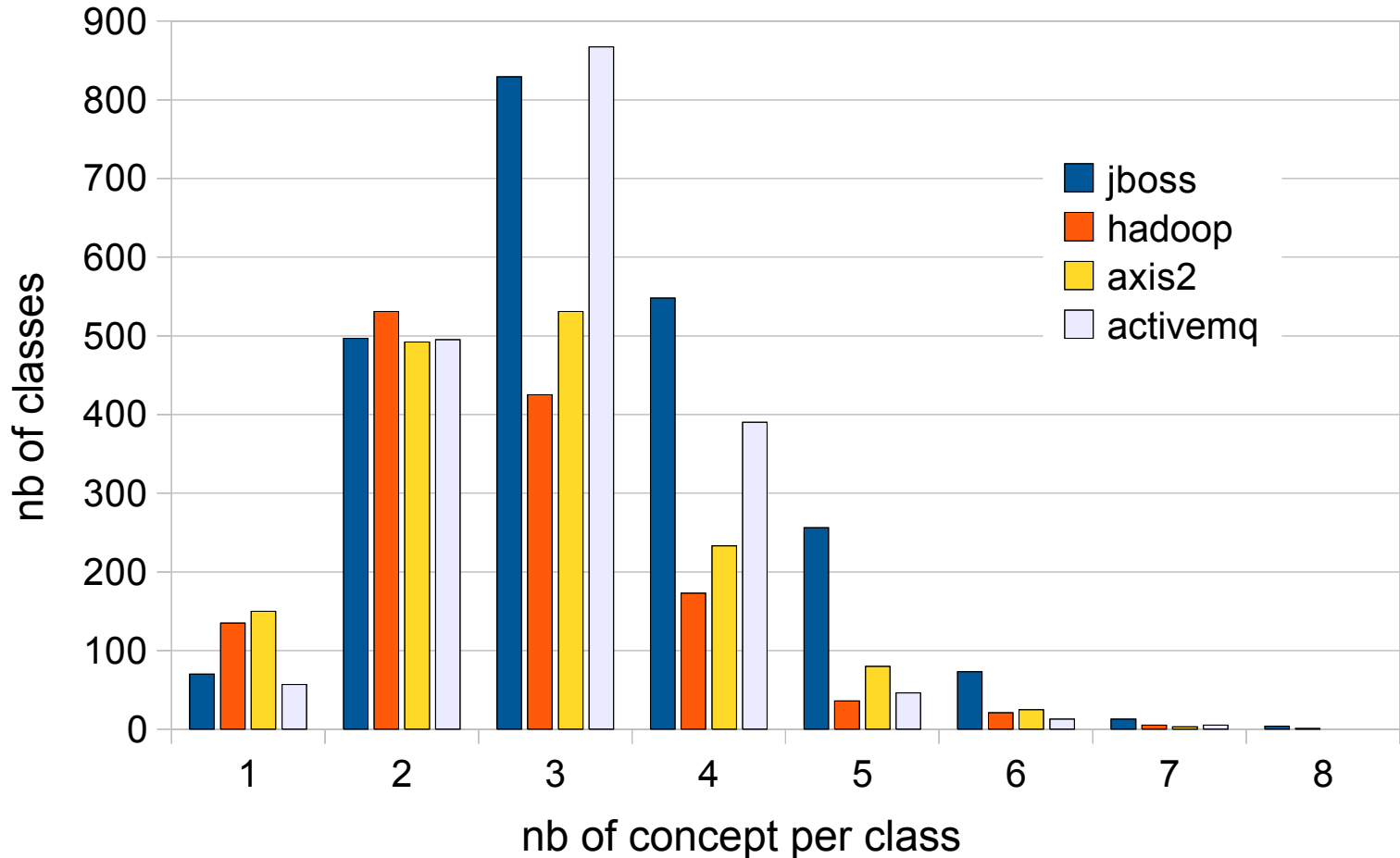
Results: Concept Frequency



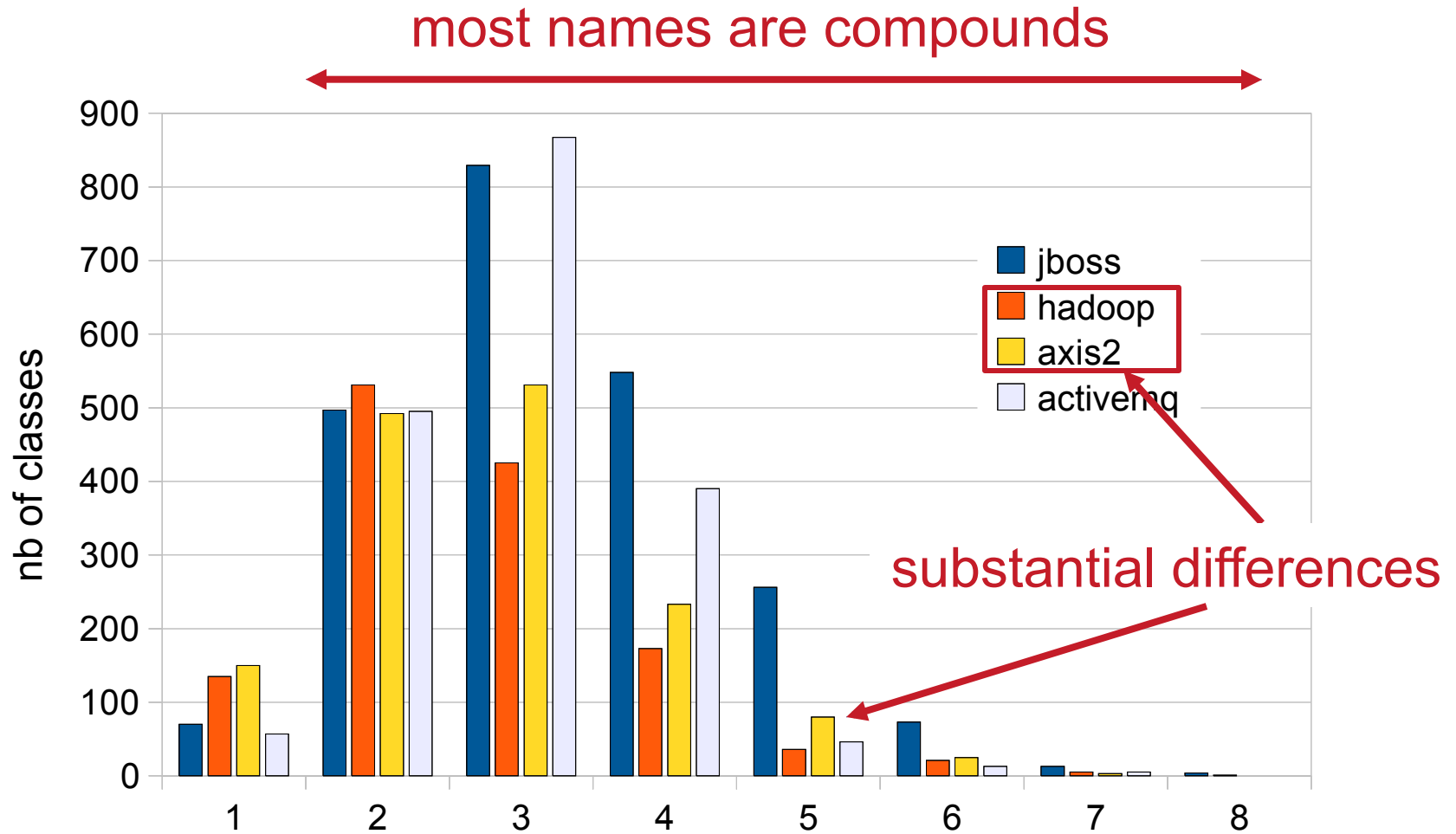
Results: Concept Frequency



Class Names Length

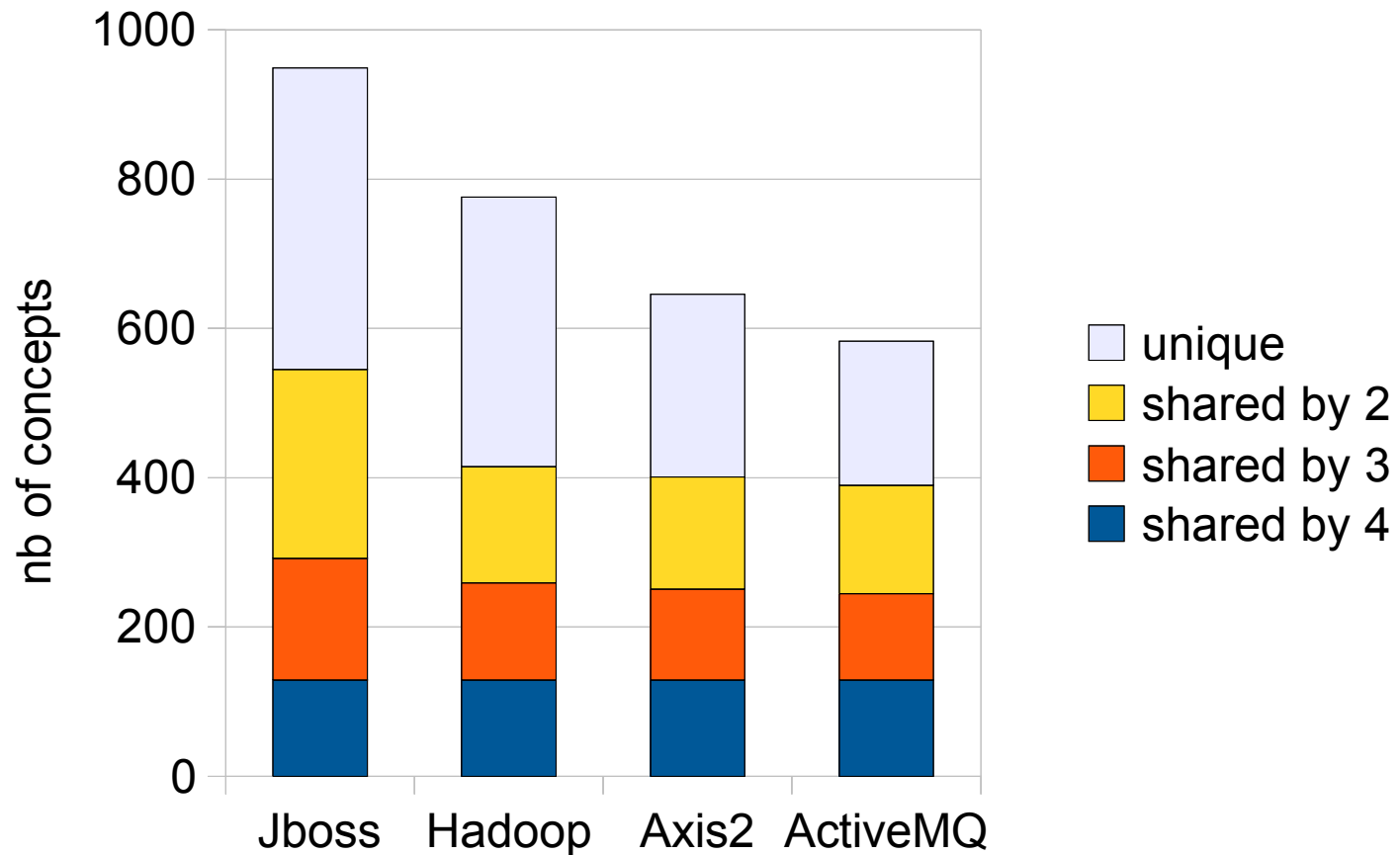


Length of Class Names



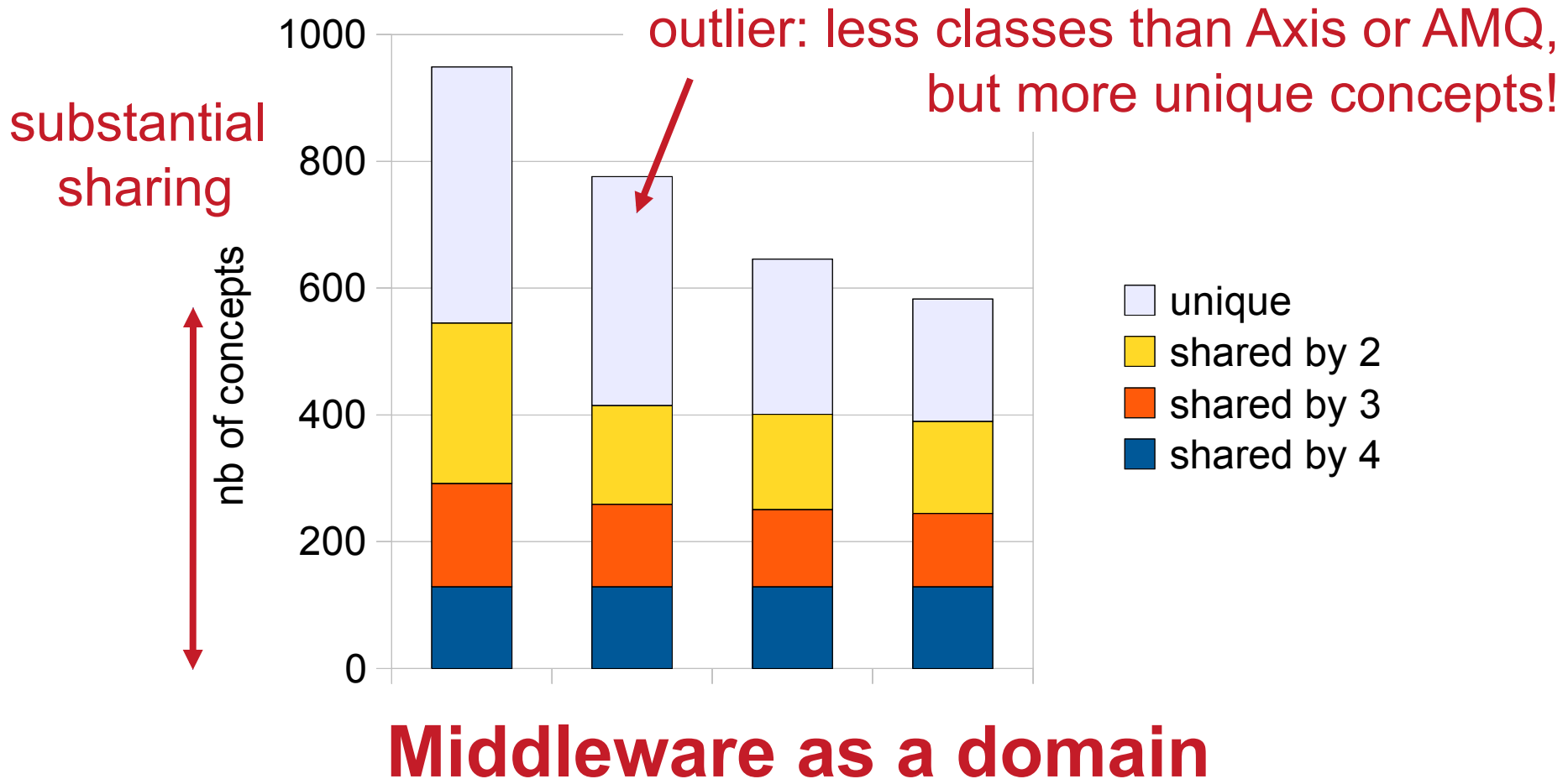
Concepts = building blocks of class names

Shared Conceptual Foundation



Shared Conceptual Foundation

- **59%** of concepts shared on average



Sharing: Common Concepts

Table 4: The top 15 concepts (in locs) shared by all projects

<i>Concept</i>	<i>locs</i>	<i>%</i>	<i>classes</i>	<i>%</i>	<i>rank</i>	<i>JBoss</i>	<i>Hadoop</i>	<i>Axis2</i>	<i>ActiveMQ</i>	<i>CV</i>
Impl	52,205	4.54%	241	3.44%	2	4.29%	0.63%	10.25%	1.93%	0.86
Service	50,948	4.43%	212	3.03%	3	5.30%	0.26%	9.30%	1.57%	0.86
<i>Data</i>	50,667	4.40%	317	4.53%	4	7.70%	2.51%	1.60%	5.19%	0.56
Message	50,013	4.35%	356	5.08%	5	1.09%	0.02%	4.25%	13.45%	1.12
Factory	42,788	3.72%	394	5.63%	6	7.34%	0.85%	2.22%	3.48%	0.70
<i>File</i>	41,795	3.63%	179	2.56%	7	0.64%	11.86%	1.33%	2.14%	1.15
Manager	37,238	3.24%	153	2.18%	8	7.46%	2.65%	0.91%	0.80%	0.92
Connection	36,856	3.20%	216	3.08%	9	5.52%	0.05%	0.43%	6.62%	0.93
<i>Context</i>	29,051	2.52%	148	2.11%	12	2.58%	0.85%	4.74%	1.48%	0.61
<i>Stream</i>	28,165	2.45%	123	1.76%	13	0.63%	3.35%	2.39%	4.11%	0.50
Abstract	24,690	2.15%	129	1.84%	16	4.45%	0.71%	1.34%	1.40%	0.74
<i>Meta</i>	23,226	2.02%	159	2.27%	18	6.30%	0.05%	0.40%	0.05%	1.56
<i>Info</i>	23,111	2.01%	193	2.76%	19	0.87%	0.63%	0.57%	6.76%	1.19
Command	21,640	1.88%	212	3.03%	20	2.76%	0.18%	0.48%	4.12%	0.87
Utils	21,003	1.83%	77	1.10%	23	0.16%	1.61%	5.34%	0.06%	1.19

Sharing: Common Concepts

- Obs1: Top concepts are **predominantly shared**

Table 4: The top 15 concepts (in locs) shared by all projects

<i>Concept</i>	<i>locs</i>	<i>%</i>	<i>classes</i>	<i>%</i>	<i>rank</i>	<i>JBoss</i>	<i>Hadoop</i>	<i>Axis2</i>	<i>ActiveMQ</i>	<i>CV</i>
Impl	52,205	4.54%	241	3.44%	2	4.29%	0.63%	10.25%	1.93%	0.86
Service	50,948	4.43%	212	3.03%	3	5.30%	0.26%	9.30%	1.57%	0.86
<i>Data</i>	50,667	4.40%	317	4.53%	4	7.70%	2.51%	1.60%	5.19%	0.56
Message	50,013	4.35%	356	5.08%	5	1.09%	0.02%	4.25%	13.45%	1.12
Factory	42,788	3.72%	394	5.63%	6	7.34%	0.85%	2.22%	3.48%	0.70
<i>File</i>	41,795	3.63%	179	2.56%	7	0.64%	11.86%	1.33%	2.14%	1.15
Manager	37,238	3.24%	153	2.18%	8	7.46%	2.65%	0.91%	0.80%	0.92
Connection	36,856	3.20%	216	3.08%	9	5.52%	0.05%	0.43%	6.62%	0.93
<i>Context</i>	29,051	2.52%	148	2.11%	12	2.58%	0.85%	4.74%	1.48%	0.61
<i>Stream</i>	28,165	2.45%	123	1.76%	13	0.63%	3.35%	2.39%	4.11%	0.50
Abstract	24,690	2.15%	129	1.84%	16	4.45%	0.71%	1.34%	1.40%	0.74
<i>Meta</i>	23,226	2.02%	159	2.27%	18	6.30%	0.05%	0.40%	0.05%	1.56
<i>Info</i>	23,111	2.01%	193	2.76%	19	0.87%	0.63%	0.57%	6.76%	1.19
Command	21,640	1.88%	212	3.03%	20	2.76%	0.18%	0.48%	4.12%	0.87
Utils	21,003	1.83%	77	1.10%	23	0.16%	1.61%	5.34%	0.06%	1.19

Key concerns are common

Sharing: Common Concepts

- Obs2: Top shared concepts **unevenly distributed**

Table 4: The top 15 concepts (in locs) shared by all projects

<i>Concept</i>	<i>locs</i>	<i>%</i>	<i>classes</i>	<i>%</i>	<i>rank</i>	<i>JBoss</i>	<i>Hadoop</i>	<i>Axis2</i>	<i>ActiveMQ</i>	<i>CV</i>
Impl	52,205	4.54%	241	3.44%	2	4.29%	0.63%	10.25%	1.93%	0.86
Service	50,948	4.43%	212	3.03%	3	5.30%	0.26%	9.30%	1.57%	0.86
<i>Data</i>	50,667	4.40%	317	4.53%	4	7.70%	2.51%	1.60%	5.19%	0.56
Message	50,013	4.35%	356	5.08%	5	1.09%	0.02%	4.25%	13.45%	1.12
Factory	42,788	3.72%	394	5.63%	6	7.34%	0.85%	2.22%	3.48%	0.70
<i>File</i>	41,795	3.63%	179	2.56%	7	0.64%	11.86%	1.33%	2.14%	1.15
Manager	37,238	3.24%	153	2.18%	8	7.46%	2.65%	0.91%	0.80%	0.92
Connection	36,856	3.20%	216	3.08%	9	5.52%	0.05%	0.43%	6.62%	0.93
<i>Context</i>	29,051	2.52%	148	2.11%	12	2.58%	0.85%	4.74%	1.48%	0.61
<i>Stream</i>	28,165	2.45%	123	1.76%	13	0.63%	3.35%	2.39%	4.11%	0.50
Abstract	24,690	2.15%	129	1.84%	16	4.45%	0.71%	1.34%	1.40%	0.74
<i>Meta</i>	23,226	2.02%	159	2.27%	18	6.30%	0.05%	0.40%	0.05%	1.56
<i>Info</i>	23,111	2.01%	193	2.76%	19	0.87%	0.63%	0.57%	6.76%	1.19
Command	21,640	1.88%	212	3.03%	20	2.76%	0.18%	0.48%	4.12%	0.87
Utils	21,003	1.83%	77	1.10%	23	0.16%	1.61%	5.34%	0.06%	1.19

Specific identity for each product

Sharing: Common Concepts

- Obs3: Large impact of **design patterns** (bold)

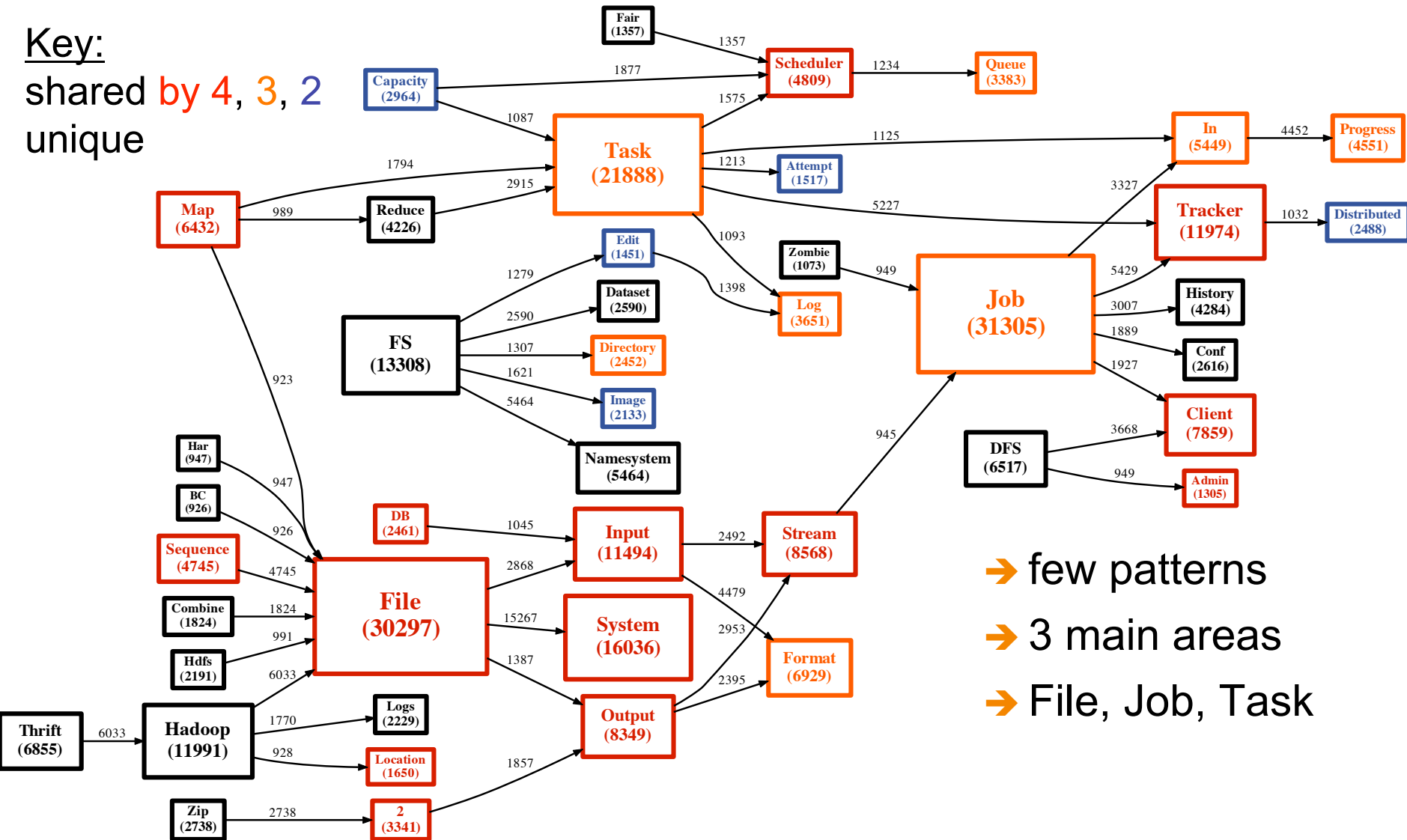
Table 4: The top 15 concepts (in locs) shared by all projects

<i>Concept</i>	<i>locs</i>	<i>%</i>	<i>classes</i>	<i>%</i>	<i>rank</i>	<i>JBoss</i>	<i>Hadoop</i>	<i>Axis2</i>	<i>ActiveMQ</i>	<i>CV</i>
Impl	52,205	4.54%	241	3.44%	2	4.29%	0.63%	10.25%	1.93%	0.86
Service	50,948	4.43%	212	3.03%	3	5.30%	0.26%	9.30%	1.57%	0.86
<i>Data</i>	50,667	4.40%	317	4.53%	4	7.70%	2.51%	1.60%	5.19%	0.56
Message	50,013	4.35%	356	5.08%	5	1.09%	0.02%	4.25%	13.45%	1.12
Factory	42,788	3.72%	394	5.63%	6	7.34%	0.85%	2.22%	3.48%	0.70
<i>File</i>	41,795	3.63%	179	2.56%	7	0.64%	11.86%	1.33%	2.14%	1.15
Manager	37,238	3.24%	153	2.18%	8	7.46%	2.65%	0.91%	0.80%	0.92
Connection	36,856	3.20%	216	3.08%	9	5.52%	0.05%	0.43%	6.62%	0.93
<i>Context</i>	29,051	2.52%	148	2.11%	12	2.58%	0.85%	4.74%	1.48%	0.61
<i>Stream</i>	28,165	2.45%	123	1.76%	13	0.63%	3.35%	2.39%	4.11%	0.50
Abstract	24,690	2.15%	129	1.84%	16	4.45%	0.71%	1.34%	1.40%	0.74
<i>Meta</i>	23,226	2.02%	159	2.27%	18	6.30%	0.05%	0.40%	0.05%	1.56
<i>Info</i>	23,111	2.01%	193	2.76%	19	0.87%	0.63%	0.57%	6.76%	1.19
Command	21,640	1.88%	212	3.03%	20	2.76%	0.18%	0.48%	4.12%	0.87
Utils	21,003	1.83%	77	1.10%	23	0.16%	1.61%	5.34%	0.06%	1.19

Code organisation main challenge?

Concept Map: Hadoop

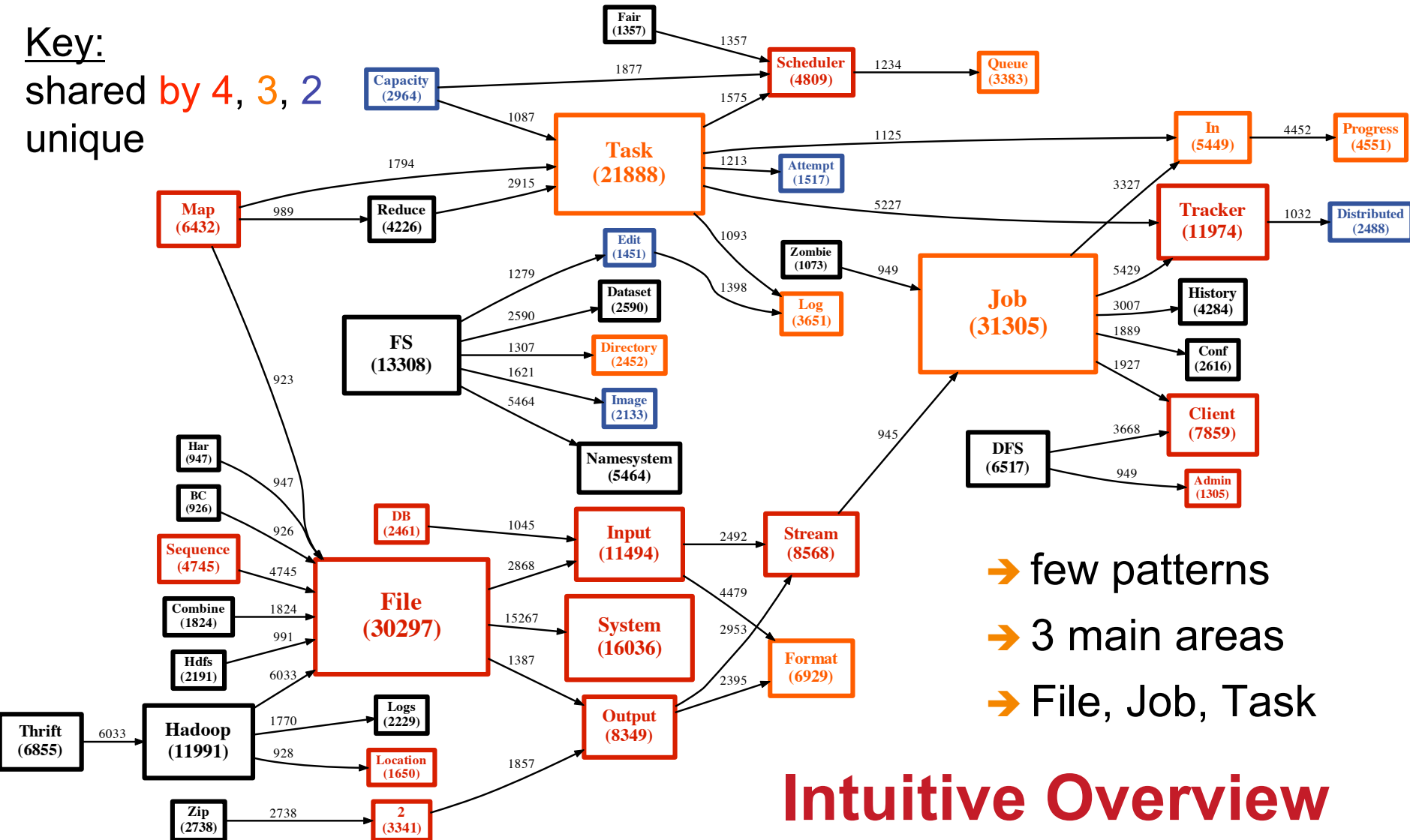
Key:
 shared by 4, 3, 2
 unique



- ➔ few patterns
- ➔ 3 main areas
- ➔ File, Job, Task

Concept Map: Hadoop

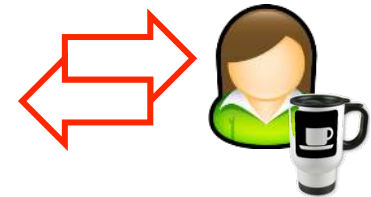
Key:
shared by 4, 3, 2
unique



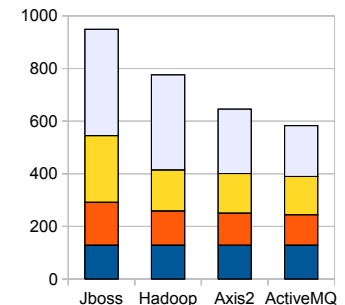
- few patterns
- 3 main areas
- File, Job, Task

Intuitive Overview

Conclusion(s)



- Middleware \neq any software
 - Human2Human aspect critical
- This study: peek into **industrial state of practice**
 - important **shared vocabulary** (59% avg)
 - key importance of **patterns** (Manager, Factory)
 - substantial **differences**: e.g. Hadoop
- Hint at “**Middleware Engineering**”
 - a field in its own right?
 - integration, composition, evolution key concerns?
- So much to look at: **exciting outlook!**



Thank you

