
体系结构设计文档

比价系统

刘璇琳 - 2016年12月22日



1.关注点

编号	关注点	类型	描述	灵活性
1	高易用性搜索	功能需求 / 质量属性	提供高易用性，兼容同义词（定义同义词表），例如搜索iphone时，可以显示苹果手机产品	基本无灵活性
2	搜索结果排序	功能需求	将符合条件的多种商品经过排序提供给用户	排序策略允许使用竞价规则调整顺序
3	屏蔽特定搜索结果	功能需求	搜索结果中可以屏蔽特定的商品，例如刚刚爆出负面新闻的商品	基本无灵活性
4	浏览器兼容	质量属性	要求用户使用时Safari、Firefox、Chrome、IE等主流浏览器均可正常显示	要求可兼容
5	爬取商品信息、价格信息	功能需求	自动从其他网站中爬取商品及其价格信息	爬取来源具有可替换性；爬取策略具有一定灵活性
6	爬取信息兼容	功能需求	要求爬取时兼容各个网站的不一致性（定义匹配格式），例如商品名称、价格单位等	实现至少2个网站的兼容；技术方案上要更加开放，例如可以兼容Amazon等英文网站
7	爬取信息更新	功能需求 / 质量属性	商品信息爬取能够定期执行，更新储存库，并且后台信息更新时不能影响在线使用	更新间隔为7天，最多为10天
8	评论商品	功能需求	用户可以查看商品信息，评论信息，并可进行评论、对评论进行回复	基本无灵活性

编号	关注点	类型	描述	灵活性
9	评论信息检查	功能需求/质量属性	用户发布评论时，检查评论中的敏感词汇（定义词汇表），发现后审查人员可以废除该评论，要可以在修改后发布	基本无灵活性，敏感词表可定制
10	水军检测	功能需求/质量属性	在用户的评论中发现敏感词汇后，自动检查该用户之前的评价词汇，分析是否是水军（敏感词汇出现大于某个数字则判定为水军），如果是则通知管理人员	检查是否为水军策略基本具有灵活性
11	商家定制候选词汇	功能需求	在用户的评论中发现敏感词汇后，如果是商品商家特殊要求的候选词汇，通知商品商家	即使不能通知商家，系统也可正常工作
12	购买商品	功能需求	用户决定购买某一商品时，可通过网站提供链接转向此商城网站进行购买	基本无灵活性
13	服务器可靠性	质量属性	服务器出现故障时要能够及时发现，并且不能影响网站访问	服务器故障发现时间在10min内；不能影响正常访问这点基本无灵活性
14	数据库可靠性	质量属性	存储数据出现故障时要能够继续工作并快速恢复	数据库出现故障不能影响访问这点基本无灵活性；数据库恢复时间小于1天
15	大存储量	质量属性	最大要能够存储1千万个商品的相关信息	基本无灵活性
16	时间效率	质量属性	保证对数据库数据的存取效率	反应时间最好在1-2s以内
17	并发性	质量属性	在高峰期时间（例如节假日），要允许500万用户同时在线使用	500万用户同时在线使用时，系统响应效率可比平时降低20%，但出现故障几率需小于1%

编号	关注点	类型	描述	灵活性
18	安全性	质量属性	进行可能的入侵检测和隔离攻击，必须实现发现短时间内频繁访问的IP并屏蔽；发现特定IP地址（例如搜索引擎Agent）的访问并屏蔽；	入侵检测规则可定制，基本无灵活性
19	可拓展性	质量属性	爬取兼容技术方案要更加开放；将来可能增加其他队访问的入侵检测规则；将来可以设计新的评论检测条件，并在满足条件的情况下通知相应人员	兼容技术方案、入侵检测规则肯定会发生变化，基本无灵活性；评论检测条件灵活性较大
20	人员	开发环境	熟悉爬虫技术、攻击检测技术的人员数量有限	灵活性不大，公司不会专门雇佣熟悉爬虫技术、攻击检测技术人员
21	团队	开发环境	项目是多团队并行开发，每个团队人数有限	灵活性不大，现状如此
22	发布计划	商业环境	项目进行分阶段研发，关于客户端的功能在后续版本中发布	第一阶段期限1年；第二阶段6个月；第三阶段2个月
23	商业目标	商业目标	系统要符合商业目标，要求系统实现竞价策略，即可按照加权的方式对某些搜索结果的排序结果进行调整	基本无灵活性，必须符合商业目标

2.体系结构需求定义

2.1 体系结构需求及其约束

优先级为1-10，数目越大优先级越高

编号	体系结构需求	优先级	约束编号	设计约束	相关约束
R1	高易用性搜索	10	C1	允许定制同义词表	阻碍C21
R2	搜索结果排序	10	C2	允许使用竞价规则定制排序策略	阻碍C21 促进C31
R3	屏蔽特定搜索结果	8	C3	允许定制屏蔽商品列表	阻碍C21
R4	浏览器兼容	8	C4	兼容Firefox、Chrome、IE、Safari最近几个版本的版本号	
R5	爬取商品信息、价格信息	10	C5	可自定义爬取来源网站	促进C11
			C6	可以替换爬取策略（深度优先，广度优先）	
R6	爬取信息兼容	10	C7	至少实现2个网站兼容	促进C25
			C8	采用更加开放的兼容技术方案，价格兼容策略、语言兼容策略	促进C25
R7	爬取信息更新	8	C9	更新间隔为7天，最多为10天	
			C10	更新信息时不能影响在线使用	
R8	评论商品	7	C11	爬取各来源网站认可度高的评论，且用户可发表评论	
R9	评论信息检查	6	C12	允许定制敏感词汇表	

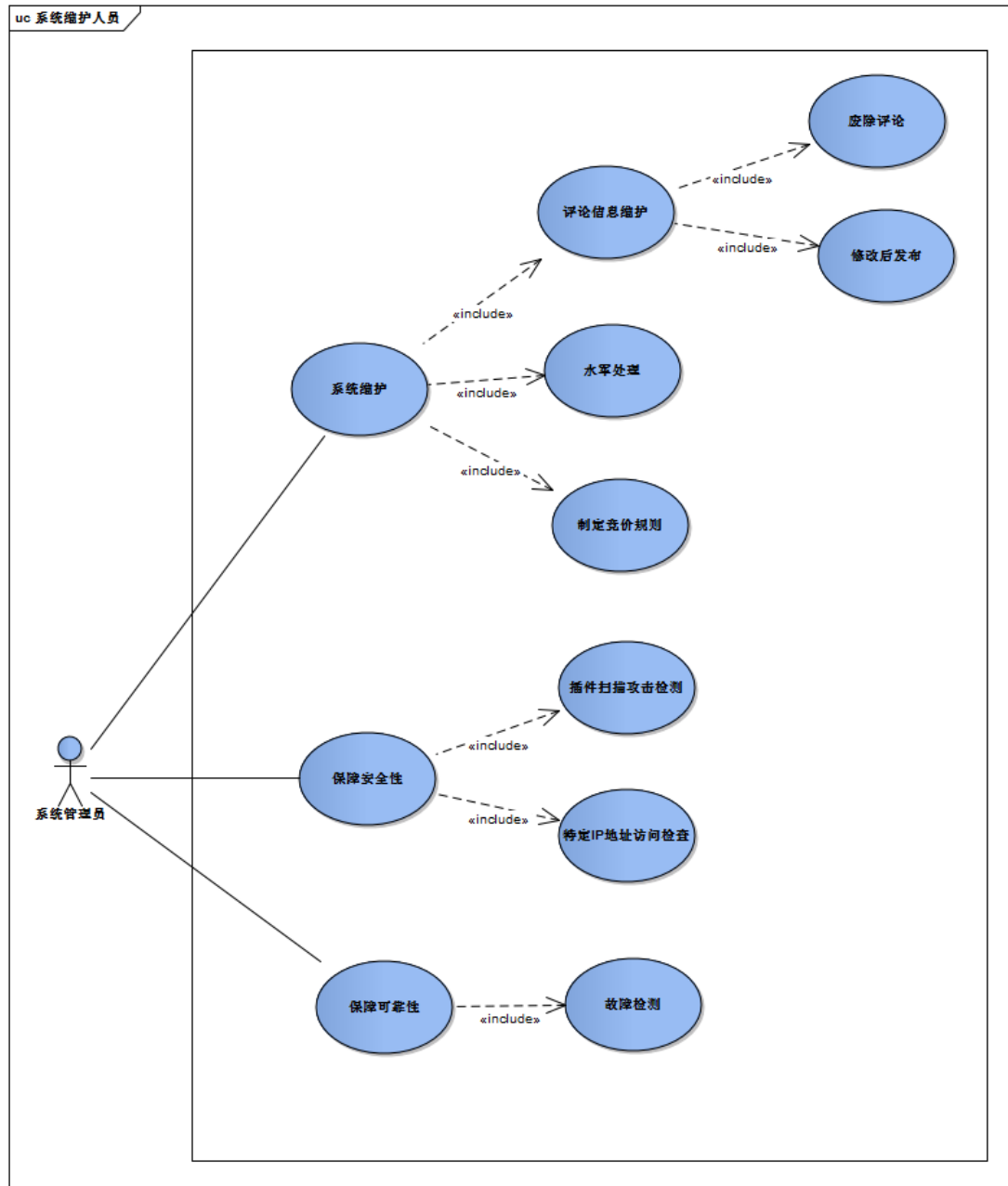
编号	体系结构需求	优先级	约束编号	设计约束	相关约束
R10	水军检测	5	C13	允许替换水军查找策略	
R11	商家定制候选词汇	5	C14	允许商家定制特殊词汇表并通知商家	
R12	购买商品	6	C15	提供转向购买商城的链接	
R13	服务器可靠性	10	C16	服务器故障应该在10min内侦测出来	促进R18
			C17	服务器故障时不能影响正常访问	
R14	数据库可靠性	10	C18	存储数据出现故障后系统能正常工作	
			C19	数据库恢复时间<12小时	
R15	大存储量	8	C20	最大能存储1千万个商品相关信息	阻碍C19阻碍C21
R16	时间效率	7	C21	反应时间最好在1s以内，1s-2s可接受	
R17	并发性	8	C22	高峰期间能允许500万用户同时在线使用	阻碍R13阻碍R14阻碍C21
R18	安全性	6	C23	发现短时间内频繁访问的IP（即插件扫描攻击）并屏蔽	促进R13促进R14
			C24	发现特定IP地址的访问并屏蔽	促进R13促进R14
R19	可拓展性	5	C25	增加一个新需兼容网站代价小于1人月	促进C7促进C8
			C26	增加一种入侵检测规则的代价小于0.25人月	

编号	体系结构需求	优先级	约束编号	设计约束	相关约束
			C27	增加新的评论检测规则代价小于0.25人月	阻碍C30
R20	人员	4	C28	信息爬取、攻击检测技术进行封装，可单独开发	
R21	团队	5	C29	采用部件技术，并且各个部件封装独立功能	促进C28
R22	发布计划	5	C30	第一阶段期限1年；第二阶段6个月；第三阶段2个月	阻碍C27
R23	商业目标	7	C31	允许定制竞价排序策略	促进C2

2.2 用例视图

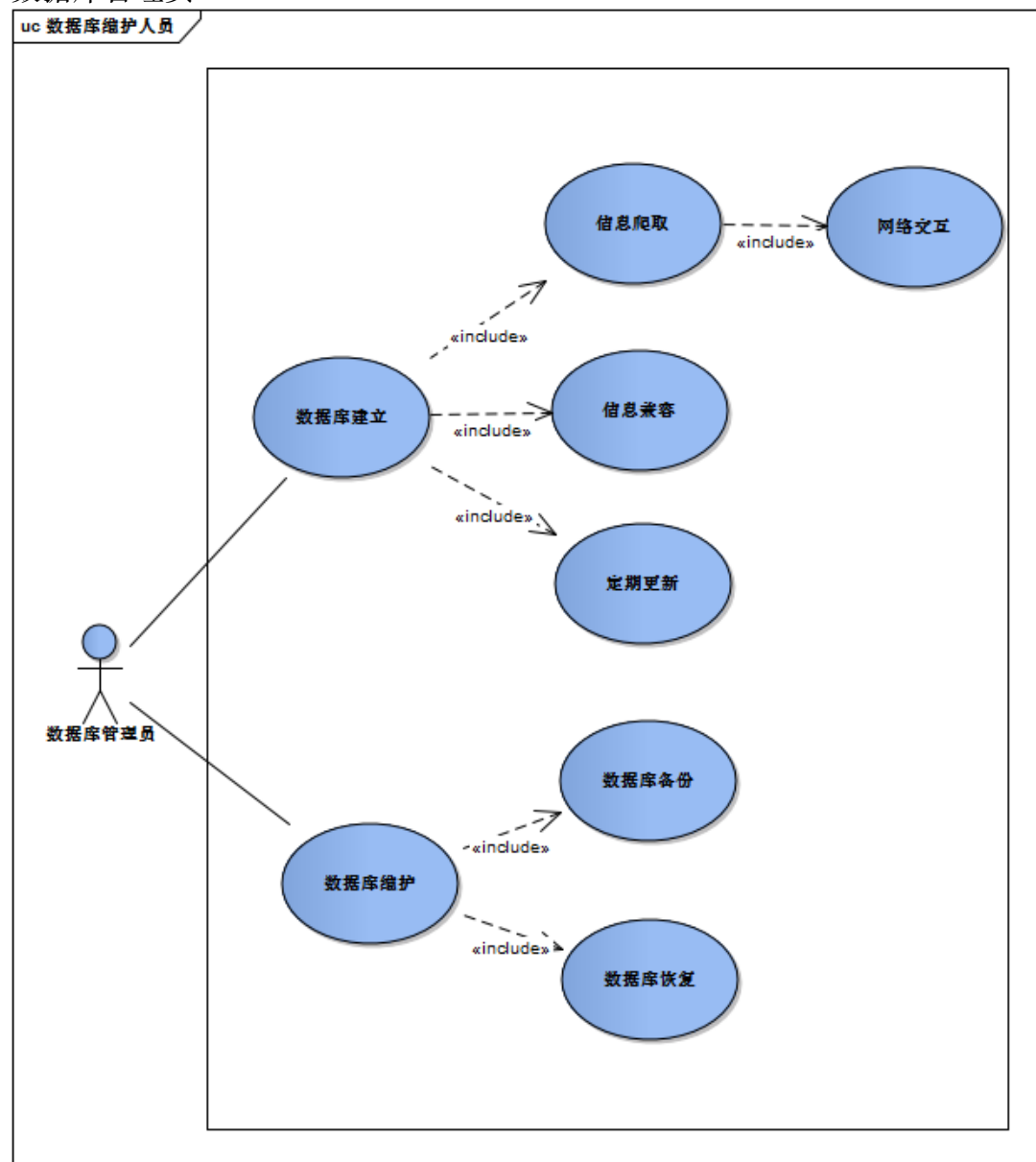
用户、商家用例图





数据库管理员

uc 数据库维护人员



2.3 场景描述

项目		内容
场景ID		S1
商业目标		用户搜索具有易用性
相关需求与设计约束		R1 高易用性搜索 C1 根据同义词表搜索
场景内容	刺激	输入“iphone”搜索请求
	刺激源	用户
	环境	正常使用时
	制品	搜索商品同义词匹配模块
	响应	根据搜索条件与同义词表进行搜索
	响应的度量	搜索结果中包含苹果手机产品

项目		内容
场景ID		S2
商业目标		搜索结果可根据竞价规则进行排序显示
相关需求与设计约束		R2 搜索结果排序 C2 允许使用竞价规则定制排序策略
场景内容	刺激	商品搜索请求
	刺激源	用户
	环境	正常使用时
	制品	搜索商品结果排序模块
	响应	根据搜索结果与竞价规则进行排序显示
	响应的度量	结果显示顺序与竞价规则误差不超过3%

项目		内容	
场景ID		S3	
商业目标		可屏蔽特定搜索结果（例如刚刚爆出负面新闻的商品）	
相关需求与设计约束		R3 屏蔽特定搜索结果	C3 可屏蔽特定搜索结果
场景内容	刺激	商品搜索请求	
	刺激源	用户	
	环境	正常使用时	
	制品	搜索商品结果屏蔽模块	
	响应	根据搜索结果与屏蔽列表进行排序显示	
	响应的度量	搜索结果中不包含屏蔽列表中商品	

项目		内容	
场景ID		S4	
商业目标		浏览器兼容	
相关需求与设计约束		R4 浏览器兼容	C4 兼容Firefox、Chrome、IE、Safari较新版本号
场景内容	刺激	网页请求：在Firefox、Chrome、IE、Safari中且版本号大于约束版本号	
	刺激源	用户	
	环境	正常使用时	
	制品	前台兼容显示模块	
	响应	返回用户请求页面	
	响应的度量	返回页面显示不正常率<1%	

项目		内容
场景ID		S5
商业目标		兼容各个网站的不一致性（定义匹配格式），例如商品名称、价格单位等
相关需求与设计约束		R6 爬取信息兼容 C7 至少实现2个网站的兼容
场景内容	刺激	开始一次信息爬取
	刺激源	数据库管理员
	环境	正常使用时
	制品	爬取信息兼容模块
	响应	从爬取源网页进行信息爬取并进行兼容操作
	响应的度量	数据库保存的价格等已爬取到信息没有兼容现象

项目		内容
场景ID		S6
商业目标		后台信息更新时不能影响在线使用
相关需求与设计约束		R7 爬取信息更新 C10 更新信息时不能影响在线使用
场景内容	刺激	启动爬取程序开始更新信息，同时请求搜索商品
	刺激源	数据库管理员，用户
	环境	正常使用时
	制品	爬取信息更新模块
	响应	爬取信息更新开始更新数据库信息
	响应的度量	搜索效率与未进行更新时相差不超过3%

项目	内容
场景ID	S7
商业目标	用户发布评论时，检查评论中的敏感词汇
相关需求与设计约束	R9 评论信息检查 C12 允许定制敏感词汇表
场景内容	刺激
	刺激源
	环境
	制品
	响应
	响应的度量

项目	内容
场景ID	S8
商业目标	发现敏感评论后自动进行水军检测
相关需求与设计约束	R10 水军检测 C13 允许替换水军查找策略
场景内容	刺激
	刺激源
	环境
	制品
	响应
	响应的度量

项目		内容
场景ID		S9
商业目标		出现的敏感词汇评论如果包含商家特殊要求的候选词汇，通知商品商家
相关需求与设计约束		R11 商家定制候选词汇 C14 允许商家定制候选词汇表并通知商家
场景内容	刺激	发现包含敏感词汇的评论
	刺激源	系统
	环境	正常使用时
	制品	候选词汇检测模块
	响应	检查评论如果包含候选词汇则通知管理员通知商品商家
	响应的度量	95%候选词汇可被检测出来

项目		内容
场景ID		S10
商业目标		服务器出现故障时要能够及时发现
相关需求与设计约束		R13 服务器可靠性 C16 服务器故障应该在10min内侦测出来
场景内容	刺激	服务器故障检测模块第三次没有听到服务器传送的状态消息
	刺激源	系统服务器
	环境	服务器故障
	制品	服务器故障侦测模块
	响应	发出故障警报
	响应的度量	故障99.9%在10min内侦测出来

项目		内容
场景ID		S11
商业目标		服务器故障时不能影响网站访问
相关需求与设计约束		R13 服务器可靠性 C17服务器故障时不能影响正常访问
场景内容	刺激	商品搜索请求
	刺激源	用户
	环境	服务器故障
	制品	服务器负载分配模块
	响应	服务器负载分配模块将此服务器相应任务分配到其他服务器
	响应的度量	网站可正常访问

项目		内容
场景ID		S12
商业目标		存储数据出现故障时要能够继续工作
相关需求与设计约束		R14 数据库可靠性 C18数据库出现故障后系统能正常继续工作
场景内容	刺激	搜索商品请求
	刺激源	用户
	环境	数据库故障时
	制品	数据库管理模块
	响应	切换使用另一个备份数据库，并返回查询信息
	响应的度量	网站可正常访问

项目	内容
场景ID	S13
商业目标	存储数据出现故障时要能够快速恢复
相关需求与设计约束	R14 数据库可靠性 C19 数据库恢复时间<1天
场景内容	刺激 数据库故障时
	刺激源 数据库
	环境 数据库故障时
	制品 数据库恢复子系统
	响应 数据库恢复子系统系统故障日志等，以便进行数据库恢复
	响应的度量 数据库恢复时间<1天

项目	内容
场景ID	S14
商业目标	大存储量的前提下保证存取效率
相关需求与设计约束	R16 时间效率 C21 反应时间最好在1s以内，1s-2s可接受
场景内容	刺激 搜索商品请求
	刺激源 用户
	环境 存储量达800万情况下
	制品 商品搜索模块
	响应 按照用户查询条件进行查询并返回结果
	响应的度量 响应时间90%在1s以内

项目		内容
场景ID		S15
商业目标		在高峰期时间（例如节假日），要允许5百万用户同时在线使用
相关需求与设计约束		R17 并发性 C22 高峰期间能允许500万用户同时在线使用
场景内容	刺激	搜索商品请求
	刺激源	用户
	环境	高峰期间（至少200万用户在线）
	制品	搜索查询模块
	响应	按照用户查询条件进行查询并返回结果
	响应的度量	高峰期为1天内服务器故障次数不超过1次

项目		内容
场景ID		S16
商业目标		必须实现短时间内频繁访问的IP（即插件扫描攻击）并屏蔽
相关需求与设计约束		R18 安全性 C23发现短时间内频繁访问的IP（即插件扫描攻击）并屏蔽
场景内容	刺激	发起网页请求
	刺激源	用户
	环境	正常使用时
	制品	插件扫描侦测模块
	响应	对客户端IP进行检测，若发现30min内访问次数超过1000次则将此IP加入IP屏蔽列表
	响应的度量	30min内访问次数超过1000次的IP被屏蔽

项目		内容
场景ID		S17
商业目标		发现特定IP地址（例如搜索引擎Agent）的访问并屏蔽
相关需求与设计约束		R18 安全性 C24发现特定IP地址的访问并屏蔽
场景内容	刺激	发起网页请求
	刺激源	用户
	环境	正常使用时
	制品	IP屏蔽模块
	响应	若此IP在IP屏蔽列表中则进行屏蔽返回Error页面
	响应的度量	IP屏蔽列表内IP不可进行访问网站

项目		内容
场景ID		S18
商业目标		技术方案上要更加开放，例如可以兼容Amazon等英文网站
相关需求与设计约束		R19 可拓展性 C25增加一个新需兼容网站代价小于1人月
场景内容	刺激	新需兼容网站
	刺激源	投资方
	环境	正常使用时
	制品	爬取信息兼容子系统
	响应	拓展爬取信息兼容子系统，可以兼容新网站
	响应的度量	拓展时间<=1人月

项目		内容	
场景ID		S19	
商业目标		将来可能还会增加其他对访问的入侵检测规则	
相关需求与设计约束		R19 可拓展性	C26增加一种入侵检测规则的代价小于0.25人月
场景内容	刺激	新的入侵检测规则	
	刺激源	投资方	
	环境	正常使用时	
	制品	入侵检测子系统	
	响应	拓展入侵检测规则，可以按照新的入侵检测规则进行检测	
	响应的度量	拓展时间<=0.25人月	

项目		内容	
场景ID		S20	
商业目标		将来可以设计新的评论检测条件，并在满足条件的情况下通知相应人员	
相关需求与设计约束		R19 可拓展性	C27增加新的评论检测规则代价小于0.25人月
场景内容	刺激	新的评论检测条件	
	刺激源	投资方	
	环境	正常使用时	
	制品	评论检测子系统	
	响应	拓展评论检测规则，可以按照新的评论检测规则进行检测	
	响应的度量	拓展时间<=0.25人月	

3.初始体系结构

3.1 来源&建立依据

系统分为主要功能为查看商品信息、搜索商品、发表评论、商品信息爬取与存储、维护评论信息、水军检测，数据库与服务器的维护。

搜索商品包含对搜索结果进行竞价排序、屏蔽特定商品；

爬取商品信息包括对信息爬取、信息解析存储、爬取信息更新。

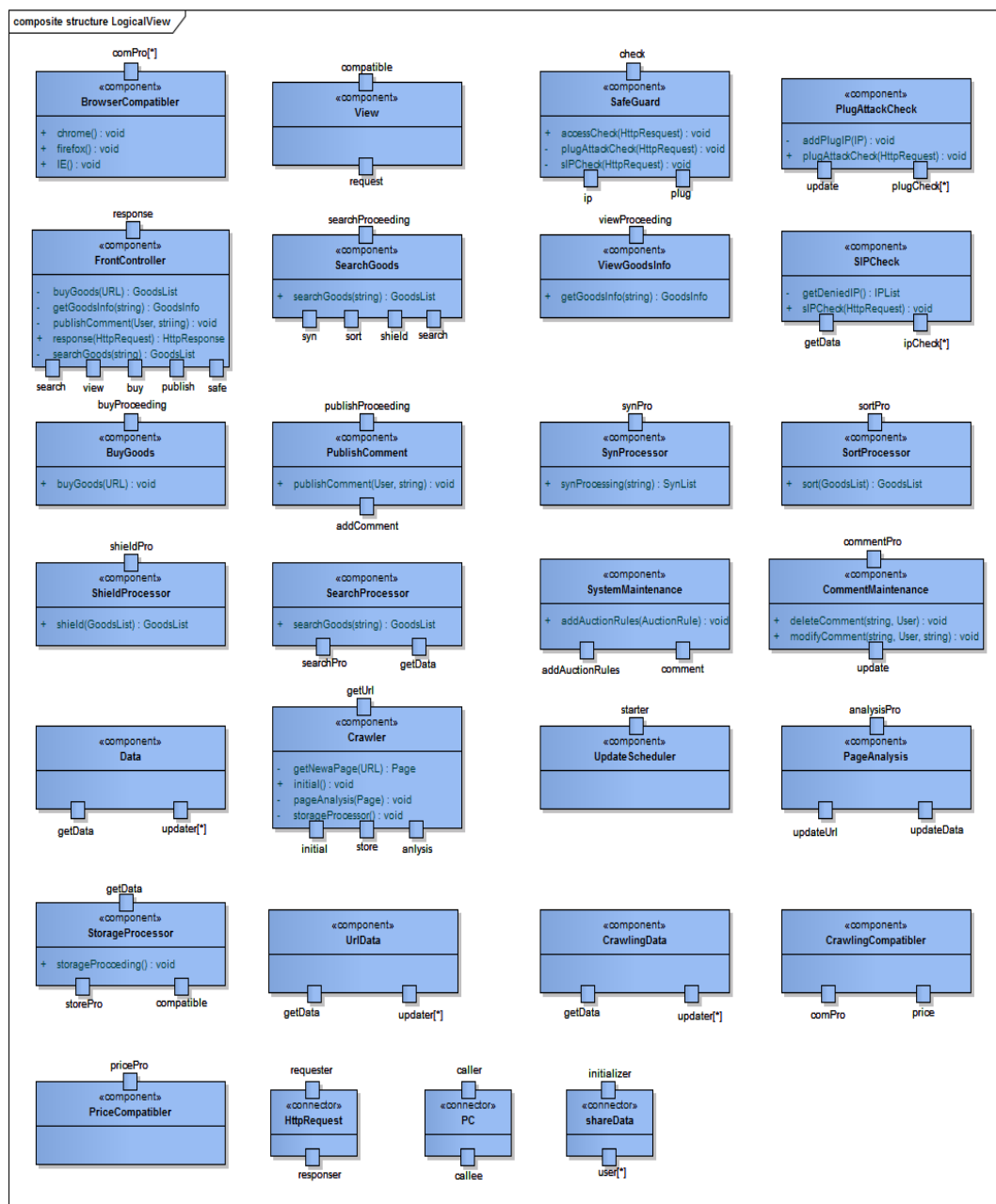
系统的维护包括攻击检测：发现短时间内频繁访问IP并屏蔽、屏蔽特定IP的访问。

因此将系统主要分为如下几个模块：搜索、查看、评论商品信息模块；爬取、存储、维护数据；入侵检测、隔离攻击；系统维护。

采用“4+1”视图描述体系结构设计，视图如下：

3.2 逻辑视图（Logical View）

比价系统的部件类型、连接件类型描述



比价系统包括25个部件类型和3个连接件类型：

- 1.部件BrowserCompatibler：主要功能是前台显示浏览器兼容。端口comPro对外提供兼容显示处理服务，可以有多个实例。
- 2.部件View的主要功能是前台显示以及向服务器发送服务请求。端口compatible负责请求浏览器兼容服务。端口request负责向服务器请求服务。
- 3.部件SafeGuard的主要功能是系统的入侵检测与隔离攻击。端口plug负责插件扫描攻击检测。端口ip负责特定IP访问检测。
- 4.部件PlugAttackCheck的主要功能是检测插件扫描攻击。端口plugCheck负责对外提供服务接口，可以有多个实例。端口update负责更新插件扫描攻击IP的共享数据。
- 5.部件FrontController的主要功能是客户端请求的分配。端口response负责响应客户端请求。端口search负责商品搜索请求。端口view负责查看商品详细信息请求。端口buy负责购买商品请求。端口publish负责发表评论请求。端口safe负责请求系统入侵检测与隔离攻击。
- 6.部件SearchGoods的主要功能是搜索商品服务。端口searchProceeding负责对外提供搜索商品服务。端口syn负责请求同义词处理服务。端口sort负责请求商品排序服务。端口shield负责请求商品屏蔽服务。端口search负责请求商品搜索处理服务。
- 7.端口ViewGoodsInfo的主要功能是查看商品详细信息。端口viewProceeding对外提供查看商品详细信息服务。
- 8.端口SIPCheck的主要功能是检测特定IP访问，例如搜索引擎。端口ipCheck负责对外提供特定IP检测服务，可以有多个实例。端口getData负责从共享数据查询信息。
- 9.部件BuyGoods的主要功能是购买商品。端口buyProceeding负责对外提供购买商品服务。
- 10.部件PublishComment的主要功能是发表评论。端口publishProceeding负责对外提供发表评论服务。端口addComment负责更新评论共享数据。
- 11.部件synProcessor的主要功能是搜索商品的同义词处理。端口synPro负责对外提供服务。
- 12.部件sortProcessor的主要功能是搜索商品的对结果进行排序。端口sortPro负责对外提供服务。

-
13. 部件shieldProcessor的主要功能是搜索商品的特定商品进行屏蔽。端口shieldPro负责对外提供服务。
14. 部件searchProcessor的主要功能是搜索商品的控制逻辑。端口searchPro负责对外提供服务。
15. 部件SystemMaintenance的主要功能是系统维护，包括增加竞价规则、评论信息的维护。端口addAuctionRules负责更新竞价信息的共享数据。端口comment负责请求评论信息维护。
16. 部件CommentMaintenance的主要功能是评论信息维护。端口commentPro负责对外提供服务。端口update负责更新评论信息的共享数据。
17. 部件Data的主要功能是维护共享数据。端口getData负责对外提供数据的使用情况。端口update负责对外提供更新数据服务，可以有多个实例。
18. 部件Crawler的主要功能是系统数据爬取。端口initial负责对外提供建立爬虫程序服务。端口store负责请求存储爬取信息。端口analysis负责请求对爬取到的信息进行分析。端口getUrl负责获取爬取来源url。
19. 部件UpdateScheduler的主要功能是开启新的爬虫程序进行数据更新。端口starter负责请求建立爬虫程序。
20. 部件PageAnalysis的主要功能是分析爬取信息，发现有用数据以及url。端口analysis负责对外提供服务。
21. 部件StorageProcessor的主要功能是存储爬取信息。端口getData负责获取分析数据来源。端口storePro负责对外提供服务。端口compatible负责请求数据兼容服务。
22. 部件UrlData的主要功能是爬取来源信息Url的共享。端口getData负责对外提供Url数据的使用情况。端口updater负责对url数据进行更新，可以有多个实例。
23. 部件CrawlingData的主要功能是爬取获取信息的共享。端口getData负责对外提供已爬取信息的使用情况。端口updater负责更新已爬取信息，可以有多个实例。
24. 部件CrawlingCompatibler的主要功能是对爬取信息进行兼容处理。端口comPro负责对外提供服务。端口price负责请求价格信息兼容处理服务。
25. 部件PriceCompatibler的主要功能是对爬取的价格信息进行兼容处理。端口pricePro负责对外提供服务。

26.连接件HttpRequest，描述HTTP协议下的远程通信机制。角色requester发送服务请求。角色responser发送服务响应。

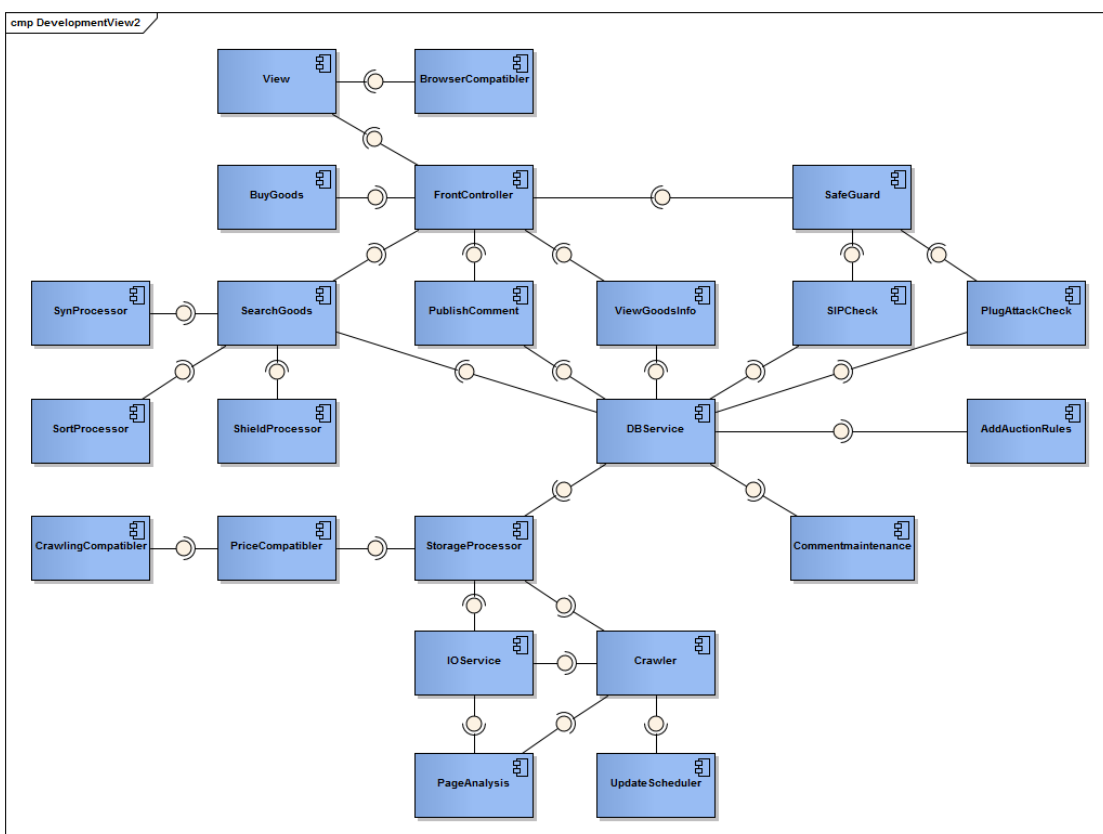
27.连接件PC（Procedural Call），描述了常见的程序调用机制。角色caller调用另一个角色callee的程序。

28.连接件ShareData，描述了共享数据机制，角色initializer将设置的数据共享给一个或多个其他角色user使用。

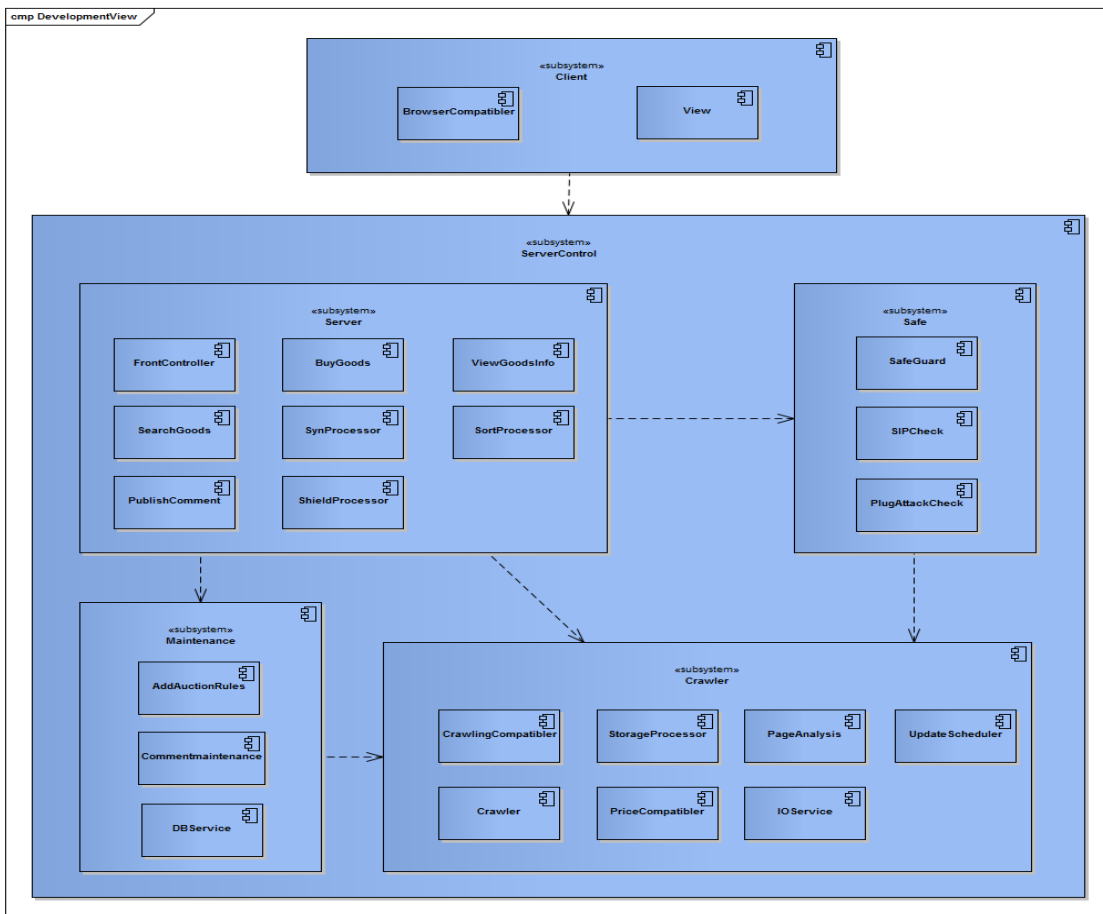
部件、连接件配置描述

3.3 开发视图

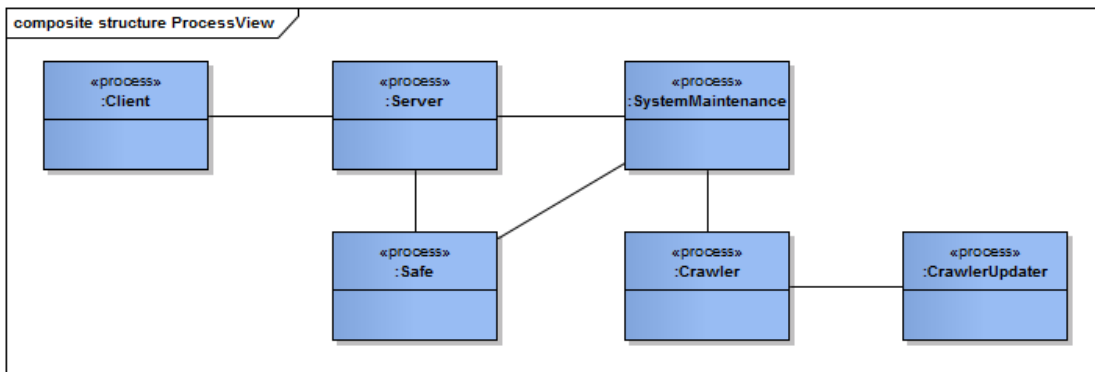
比价系统的模块组织



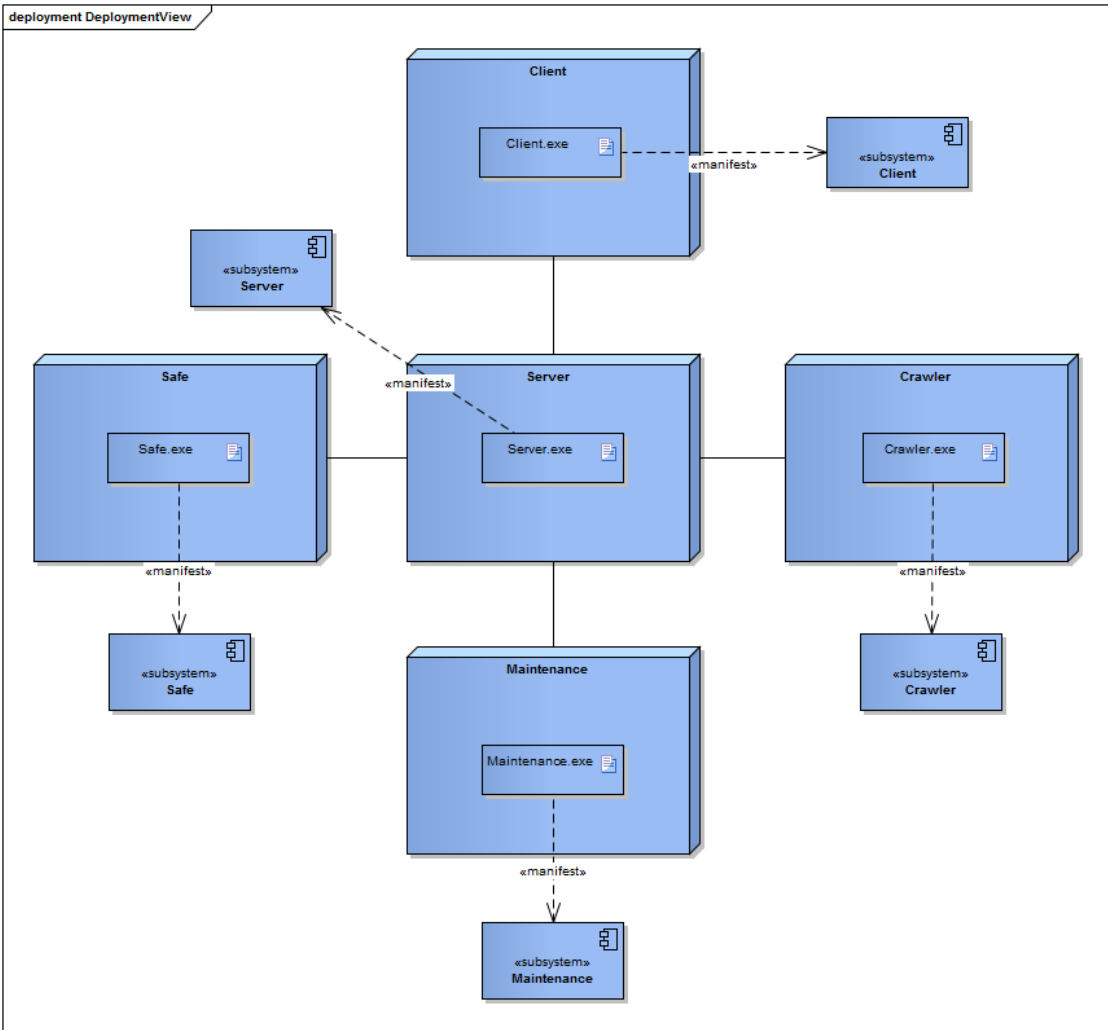
比价系统的子系统组织



3.4 进程视图



3.5 部署视图



4.设计决策

4.1 设计决策1

R1 高易用性搜索 C1 允许定制同义词表	R9 评论信息检查 C12允许定制敏感词汇表
R3 屏蔽特定搜索结果 C3 允许定制屏蔽商品列表	R11 商家定制候选词汇 C14 允许商家定制特殊词汇表
方案：同义词表、屏蔽商品列表、敏感词汇表、候选词汇表存储于数据库中； 增加同义词字典，屏蔽列表字典，敏感词字典，候选词汇表字典，封装对数据访问与更新，使得数据与逻辑相分离	
详细设计约束：封装数据访问细节	

4.2 设计决策2

R2 搜索结果排序

C2 允许使用竞价规则定制排序策略

R23 商业目标

C31 允许定制竞价排序规则

方案：封装竞价排序规则，Strategy模式

影响：逻辑视图、开发视图

详细设计约束：隐藏竞价排序规则细节

4.3 设计决策3

R5 爬取商品信息、价格信息

C5 可自定义爬取来源网站

C6 可以替换爬取策略，例如深度优先、广度优先

方案：封装页面爬取算法以及起始页面，Strategy模式

影响：逻辑视图、开发视图

详细设计约束：隐藏页面爬取算法

4.4 设计决策4

R4 浏览器兼容

方案：实现兼容策略

影响：在初始方案中已实现

4.5 设计决策5

R6 爬取信息兼容

C7 至少实现2个网站兼容

C8采用更加开放的兼容技术方案，价格兼容策略、语言兼容策略单独抽取

R19 可拓展性

C25增加一个新需兼容网站代价小于1人月

方案：封装每个网站的兼容技术方案，并封装价格兼容方案、语言兼容方案

理由：封装价格、语言兼容方案利于复用；封装每个网站兼容技术方案符合OCP原则

影响：逻辑视图、开发视图

详细设计约束：隐藏兼容技术方案细节

4.6 设计决策6

方案：采取分幅更新，每次将更新幅从数据库取出以供查询，更新数据库。

影响：逻辑视图、开发视图

4.7 设计决策7

R9 评论信息检查

C12 允许定制敏感词汇表

R10 水军检测

C13 允许替换水军查找策略

R11 商家定制候选词汇

C14允许商家定制特殊词汇表并通知商家

方案：采用事件监听机制，对评论信息进行检测，并定制不同处理策略，符合OCP原则

影响：逻辑视图、开发视图、进程视图、部署视图

详细设计约束：事件监听机制

4.8 设计决策8

R19 可拓展性

C27增加新的评论检测规则代价小于0.25人月

方案：封装评论检测规则，Strategy模式

影响：逻辑视图、开发视图

详细设计约束：隐藏评论检测规则细节

4.9 设计决策9

R13 服务器可靠性

C16 服务器故障应该在10min内侦测出来

方案：使用Ping/Echo方法检测服务器故障

影响：逻辑视图、开发视图、进程视图、部署视图

详细设计约束：Ping/Echo规则

4.10 设计决策10

R13 服务器可靠性

R17 服务器故障时不能影响正常访问

R17 并发性

C22 高峰期间能允许500万用户同时在线使用

R16 时间效率

C21反应时间最好在1s以内，1s-2s可接受

方案：使用多个处理器进行处理，并且增加负载均衡分配器

影响：逻辑视图、开发视图、进程视图、部署视图

详细设计约束：分配器允许动态增加、删除处理器

新的体系结构约束CA1：一致性更新与Cluster访问

4.11 设计决策11

R14 数据库可靠性

C18存储数据出现故障后系统能正常继续工作

C19数据库恢复时间<1天

R15 大存储量

C20最大能存储1千万个商品相关信息

R16 时间效率

C21反应时间最好在1s以内，1s-2s可接受

方案：使用分布式存储，并且增加数据库访问分配器以对数据库进行管理；
利用数据库热备份有利于容灾；
分布式存储有利于容灾

影响：逻辑视图、开发视图、进程视图、部署视图

详细设计约束：数据库故障时，分配器能够随时启动备份数据库

4.12 设计决策12

R18 安全性

C23发现短时间内频繁访问的IP（即插件扫描攻击）并屏蔽

C24发现特定IP地址的访问并屏蔽

R19 可拓展性

C26增加一种入侵检测规则的代价小于0.25人月

方案：封装入侵检测规则，并封装插件扫描攻击、屏蔽特定IP模块，信息隐藏，增加controller对入侵检测规则进行管理

影响：逻辑视图，开发视图

详细设计约束：隐藏入侵检测规则细节

4.13 设计决策13

R22 发布计划

C30第一阶段期限1年；第二阶段6个月；第三阶段2个月

方案：封装模块，各小组独立平行开发

使用中间件平台作为支撑，便利RPC机制开发

影响：开发视图、进程视图

详细设计约束：中间件平台接口；进程的方法调用接口

4.14 设计决策14

R13 服务器可靠性

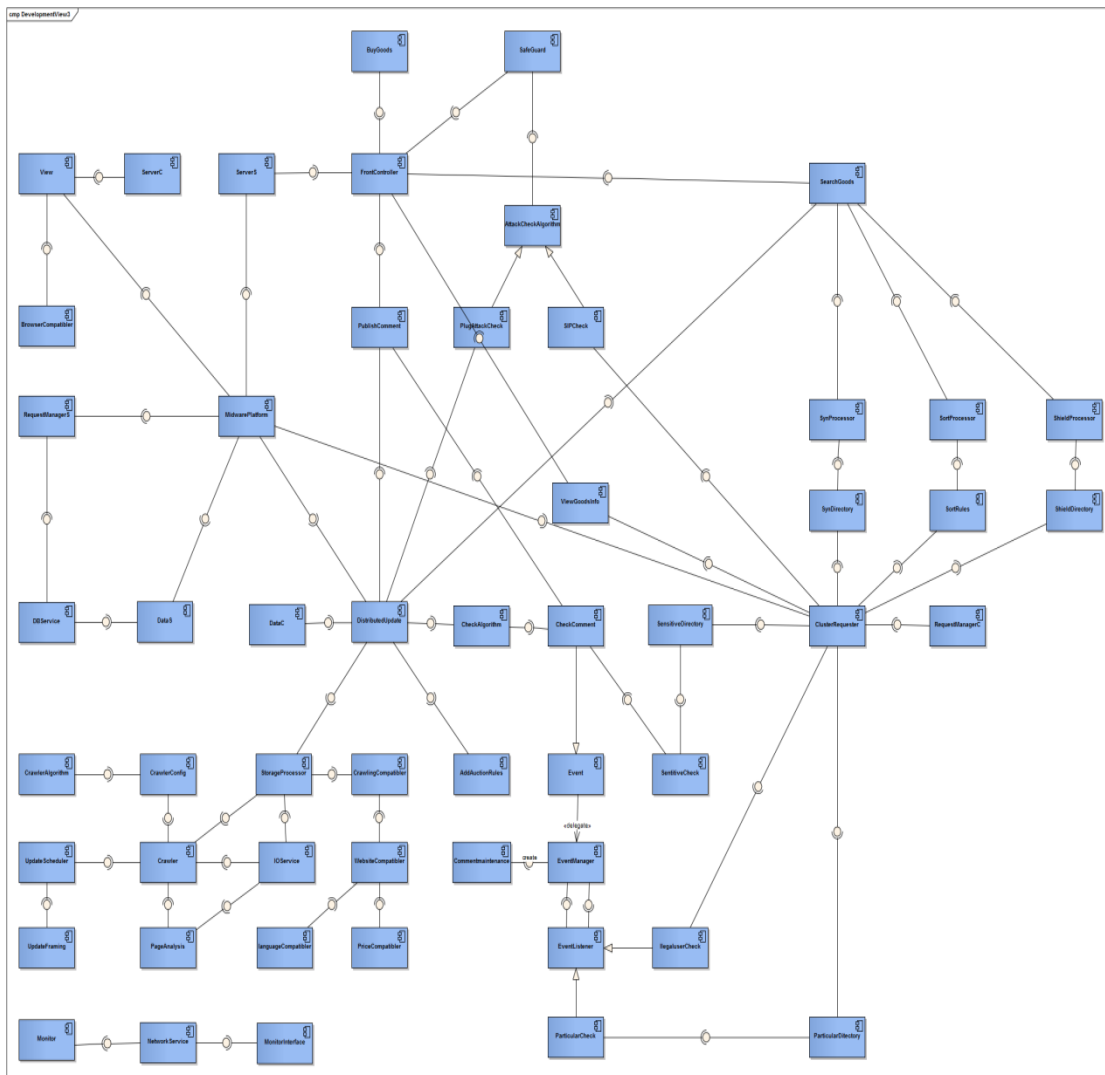
CA1 一致性更新与Cluster访问

方案：利用中间件平台简化一致性更新与Cluster访问

影响：开发视图

[illegible]

5.2 开发视图



逻辑视图与开发视图之间的映射关系描述

部件/连接件	实现模块/构件
View	View
BrowserCompatibler	BrowserCompatibler
FrontController	FrontController、MonitorInterface
Monitor	Monitor
PC	无
HttpRequest	socket
DistributeRPC	DistributeUpdate, Middleware Platform
ClusterRPC	ClusterRequestor, Middleware Platform
PingEcho	Network Service
ShareData	IOService, DBService
Generation	(generation)
EventManager	EventManger
Event	Event
EventListener	EventListener

开发视图接口描述

ID: 1		构件: View, ServerC	
供给接口			
ServerC			
接口ID	功能描述	语法	void response(HttpRequest request)
		前置条件	Request符合Http协议
		后置条件	无
		不变量	无
需求接口			
View.request(HttpRequest)			

ID: 2		构件: ServerS, FrontController	
供给接口			
FrontController			
接口ID	功能描述	语法	void response(HttpRequest request)
		前置条件	Request符合Http协议
		后置条件	无
		不变量	无
需求接口			
ServerS.response(HttpRequest)			

ID: 3		构件: FrontController,SearchGoods	
供给接口			
SearchGoods			
接口ID	功能描述	语法	GoodsList searchGoods(string description)
		前置条件	description为中文 或英文
		后置条件	无
		不变量	无
需求接口			
FrontController.searchGoods(String)			

ID: 4	构件: FrontController,ViewGoodsInfo
供给接口	
ViewGoodsInfo	

接口ID	功能描述	语法	GoodsInfo getGoodsInfo(stirng goodsID)
		前置条件	goodsID为合法ID
		后置条件	无
		不变量	无
需求接口			
FrontController.viewGoodsInfo(Strings)			

ID: 5		构件: FrontController,BuyGoods	
供给接口			
BuyGoods			
接口ID	功能描述	语法	void buyGoods(URL url)
		前置条件	url合理
		后置条件	跳转往url指向网站
		不变量	无
需求接口			
FrontController.buyGoods(URL)			

ID: 6		构件: FrontController, PublishComment	
供给接口			
PublishComment			
接口ID	功能描述	语法	void publishComment(User user,String content)
		前置条件	User为合法用户

需求接口	后置条件		User.comments.add (content)
	不变量		无
	FrontController.publishComment(String,User)		

ID: 7		构件: SearchGoods,SynProcessor	
供给接口			
SynProcessor			
接口ID	功能描述	语法	List synProcessing(String content)
		前置条件	同义词列表已存在
		后置条件	无
		不变量	无
需求接口			
SearchGoods.getSynList(String)			

ID: 8		构件: SearchGoods,ShieldProcessor	
供给接口			
ShieldProcessor			
接口ID	功能描述	语法	GoodsList shield(GoodsList list)
		前置条件	GoodsList中goods 为合法商品，在数 据库中保存
		后置条件	屏蔽列表已存在
		不变量	无
需求接口			
SearchGoods.shield(GoodsList)			

ID： 9		构件： SearchGoods,SortProcessor	
供给接口			
SortProcessor			
接口ID	功能描述	语法	GoodsList sort(GoodsList list)
		前置条件	GoodsList中goods 为合法商品， 在数 据库中保存
		后置条件	无
		不变量	无
需求接口			
SearchGoods.sort(GoodsList)			

ID：10		构件：SynProcessor,SynDirectory	
供给接口			
SynDirectory			
接口ID	功能描述	语法	List getSynList(String content)
		前置条件	同义词列表已存在
		后置条件	无
		不变量	无
需求接口			
SynProcessor.getSynList(String)			

ID: 11			构件: SortProcessor,SortRules
供给接口			
SortRules			
接口ID	功能描述	语法	SortRuleType getSortRule()

	前置条件	已声明 SortRuleType
	后置条件	无
	不变量	无
需求接口		
SortProcessor.getSortRule()		

ID: 12		构件: ShieldProcessor,ShieldDirectory	
供给接口			
ShieldDirectory			
接口ID	功能描述	语法	List getShieldList()
		前置条件	屏蔽列表已存在
		后置条件	无
		不变量	无
需求接口			
ShieldProcessor.gerShiledList()			

ID: 13		构件: FrontController,SafeGuard	
供给接口			
SafeGuard			
接口ID	功能描述	语法	access(IPType ip)
		前置条件	Ip为合法ip地址
		后置条件	无
		不变量	无
需求接口			
FrontController.access(IPType)			

ID: 14		构件: SafeGuard,AttackCheckAlgorithm	
供给接口			
AttackCheckAlgorithm			

接口ID	功能描述	语法	access(IPType ip)
		前置条件	Ip为合法ip地址
		后置条件	无
		不变量	无
需求接口			
SafeGuard.access(IPType)			

ID： 15		构件： AttackCheckAlgorithm,SIPCheck	
供给接口			
SIPCheck			
接口ID	功能描述	语法	boolean isSpecialIP(IPType ip)
		前置条件	Ip为合法ip地址， 特殊ip地址列表已存在
		后置条件	无
		不变量	无
需求接口			
AttackCheckAlgorithm.isSpecialIP(IPType)			

ID: 16		构件: AttackCheckAlgorithm,PlugAttackCheck	
供给接口			
PlugAttackCheck			
接口ID	功能描述	语法	boolean checkAttack(IPType ip)
		前置条件	Ip为合法ip地址
		后置条件	无
		不变量	无
需求接口			
AttackCheckAlgorithm.checkAttack(IPType)			

ID: 17		构件: UpdateScheduler,UpdateFraming	
供给接口			
UpdateFraming			
接口ID	功能描述	语法	void framing(FrameContent content)
		前置条件	数据库已建立且正常运行
		后置条件	无
		不变量	content=this.framContent; UpdateFraming中 游标cursor指向下一个
需求接口			
UpdateScheduler.framing(FrameContent)			

ID: 18		构件: UpdateShedular,Crawler	
供给接口			
Crawler			
接口ID	功能描述	语法	void initial()
		前置条件	updateShedular经过间隔period做出更新决定
		后置条件	updateShedular的计时器清零
		不变量	无
需求接口			
UpdateShedular.start(Time)			

ID: 19	构件: Crawler,CrawlerConfig
供给接口	
CrawlerConfig	

接口ID	功能描述	语法	configCrawler(CrawlerRuleType type,SourceList list)
		前置条件	Type为枚举类型,list中存放爬取网页来源url
		后置条件	CrawlerTuleType.type=type,SourList.list=list
		不变量	无
需求接口			
Crawler.config(CrawlerRuleType)			

ID: 20		构件: CrawlerConfig,CrawlerAlgorithm	
供给接口			
CrawlerAlgorithm			
接口ID	功能描述	语法	CrawlerRule getCrawlerAlgorithm(CrawlerRuleType type)
		前置条件	Type为枚举类型，CrawlerRule已设置完毕
		后置条件	无
		不变量	无
需求接口			
CrawlerConfig. getCrawlerAlgorithm(CrawlerRuleType)			

ID: 21		构件: Crawler,IOService	
供给接口			
IOService			
接口ID	功能描述	语法	Url getNextUrl()
		前置条件	IOService中存在urlList

	后置条件	IOService中 UrlData游标cursor 后移一位
	不变量	无
需求接口		
Crawler.getNextUrl()		

ID: 22		构件: Crawler,PageAnalysis	
供给接口			
PageAnalysis			
接口ID	功能描述	语法	Analyze(WebPage page)
		前置条件	page已获得
		后置条件	Page已经过处理存入文件
		不变量	无
需求接口			
Crawler.analyzeRequest(WebPage)			

ID： 23		构件： PageAnalysis,IOService	
供给接口			
IOService			
接口ID	功能描述	语法	void addUrl(Url url)
		前置条件	url为合法url地址
		后置条件	urlData.add(url)
		不变量	无
接口ID	功能描述	语法	void setWebPageContent(WebPageContent content)
		前置条件	WebPageContent为已经解析好的wen网页内容

ID： 24		构件： Crawler,StorageProcessor	
供给接口			
StorageProcessor			
接口ID	功能描述	语法	void store()
		前置条件	pageAnalysis已经将解析好的页面存入文件
		后置条件	从解析好的文件中获取到有效信息，并存入数据库
		不变量	无
需求接口			
Crawler. store()			

ID: 25		构件: StorageProcessor,CrawlingCompatibler	
供给接口			
CrawlingCompatibler			
接口ID	功能描述	语法	void compatible(WebPageContent content)
		前置条件	StorageProcessor已经从文件系统中取出待保存页面
		后置条件	将商品信息以统一格式保存入数据库
		不变量	无
需求接口			
StorageProcessor.compatible(WebPageContent)			

ID: 26		构件: StorageProcessor,IOService	
供给接口			

IOService			
接口ID	功能描述	语法	WebPageContent getContent()
		前置条件	pageAnalysis已经 将解析好的页面存 入文件
		后置条件	将页面信息进行兼 容保存
		不变量	无
需求接口			
StorageProcessor .getContent()			

ID: 27		构件: CrawlingCompatibler,WebsiteCompatibler	
供给接口			
WebsiteCompatibler			
接口ID	功能描述	语法	void compatible(WebPageContent content)
		前置条件	StorageProcessor已经从文件系统中取出待保存页面
		后置条件	商品信息以统一形式存入数据库
		不变量	无
需求接口			
CrawlingCompatibler.compatible(WebPageContent)			

ID: 28	构件: WebsiteCompatibler,PriceCompatibler
供给接口	
PriceCompatibler	

接口ID	功能描述	语法	Price setPrice(String content)
需求接口		前置条件	content为已从页面中解析好的价格信息
		后置条件	利用content中价格信息构造统一形式的Price
		不变量	无
		WebsiteCompatibler.setPrice(String)	

ID: 29		构件: WebsiteCompatibler,LanguageCompatibler	
供给接口			
LanguageCompatibler			
接口ID	功能描述	语法	String setLanguage(String content)
		前置条件	content为从页面中 解析出来的某种语 言的信息
		后置条件	将content转化为另 一种语言信息
		不变量	无
需求接口			
WebsiteCompatibler. setLanguage(String)			

ID: 30		构件: CommentMaintenance,EventManager	
供给接口			
EventManager			
接口ID	功能描述	语法	void initial()

	前置条件	commentMaintenance已经建立
	后置条件	eventManager被创建
	不变量	无
需求接口		
CommentMaintenance.createEventManager()		

ID: 31		构件: EventManager,Event	
供给接口			
Event			
接口ID	功能描述	语法	void initial()
		前置条件	EventManager已经创建
		后置条件	Event加入到eventManager管理的事件队列中
		不变量	无
需求接口			
EventManager.createEvent()			

ID: 32		构件：EventListener, EventManager	
供给接口			
EventManager			
接口ID	功能描述	语法	void register(Event source, EventListener listener)
		前置条件	Event已经被创建
		后置条件	EventManager.addToEvent(source,listener)
		不变量	无

需求接口			
EventListener.register(Event)			

ID: 33		构件: EventManager, EventListener	
供给接口			
EventListener			
接口ID	功能描述	语法	void perform(Event event)
		前置条件	event发生
		后置条件	处理此event发生之后的情况
		不变量	无
需求接口			
EventManager.inform(Event,EventListenerList)			

ID: 34		构件: CheckComment,CheckAlgorithm	
供给接口			
CheckAlgorithm			
接口ID	功能描述	语法	void setAlgorithm(AlgorithmType type)
		前置条件	Type为枚举类型
		后置条件	Algorithm.setType(type)
		不变量	无
需求接口			
CheckComment. setAlgorithm(AlgorithmType)			

ID: 35		构件: CheckComment,SensitiveCheck	
供给接口			
SensitiveCheck			

接口ID	功能描述	语法	void sensitiveCheck(List contentList)
		前置条件	contentList中内容 为用户发表的评论 信息
		后置条件	无
		不变量	无
需求接口			
CheckComment.check()			

ID: 36		构件: SensitiveCheck,SensitiveDirectory	
供给接口			
SensitiveDirectory			
接口ID	功能描述	语法	boolean isSensitive(String content)
		前置条件	content为用户发表的 评论信息，敏感 词汇表已经创建
		后置条件	无
		不变量	无
需求接口			
SensitiveCheck. isSensitive(String)			

ID: 37			构件: PaticularCheck,ParticularDirectory
供给接口			
ParticularDirectory			
接口ID	功能描述	语法	boolean isParticular(String content)

	前置条件	content为已被检测出包含有敏感词的comment信息
	后置条件	无
	不变量	无
需求接口		
PaticularCheck.check()		

ID：38		构件： SensitiveDirectory,ClusterRequester	
供给接口			
ClusterRequester			
接口ID	功能描述	语法	List getSensitiveList()
		前置条件	敏感词列表已经创建
		后置条件	无
		不变量	无
需求接口			
SensitiveDirectory.getList()			

ID：39		构件：IllegaluserCheck, ClusterRequester	
供给接口			
ClusterRequester			
接口ID	功能描述	语法	UserInfo getUserInfo(String userID)
		前置条件	userID为合法ID
		后置条件	检查用户的评论包 含敏感信息的数量
		不变量	无
需求接口			

IllegaluserCheck.getUserCommentHistory()

ID: 40

构件: ParticularDirectory,
ClusterRequester

供给接口

ClusterRequester

接口ID	功能描述	语法	List getParticularList()
		前置条件	特殊词汇表已经被创建
		后置条件	无
		不变量	无

需求接口

ParticularDirectory.getList()

ID: 41

构件: SynDirectory, ClusterRequester

供给接口

ClusterRequester

接口ID	功能描述	语法	List getSynList()
		前置条件	同义词表已经被创建
		后置条件	无
		不变量	无

需求接口

SynDirectory.getList()

ID: 42

构件: SortRules, ClusterRequester

供给接口

ClusterRequester			
接口ID	功能描述	语法	String getSortRules()
		前置条件	竞价规则已经被创建
		后置条件	将string信息解析为SortRule
		不变量	无
需求接口			
SortRules.getRules()			

ID：43		构件：ShieldDirectory, ClusterRequester	
供给接口			
ClusterRequester			
接口ID	功能描述	语法	List getShiedList()
		前置条件	商品屏蔽列表已经被创建
		后置条件	无
		不变量	无
需求接口			
ShieldDirectory.getList()			

ID: 44		构件: addAuctionRule,DistributedUpdate	
供给接口			
DistributedUpdate			
接口ID	功能描述	语法	void update(SortRule rule)
		前置条件	rule为已经解析好的 信息

	后置条件	数据库中保存竞价规则信息，所有数据库同时更新
	不变量	无
需求接口		
addAuctionRule.add(SortRule)		

ID: 45		构件: StorageProcessor, DistributedUpdate	
供给接口			
DistributedUpdate			
接口ID	功能描述	语法	void update(String content)
		前置条件	content为爬取程序爬取到的商品信息，且已经被兼容处理过，形式统一
		后置条件	数据库中保存content商品信息，所有数据库同时更新
		不变量	无
需求接口			
StorageProcessor.store()			

ID: 46		构件: PublishComment, DistributedUpdate	
供给接口			
DistributedUpdate			
接口ID	功能描述	语法	void addComment(Comment comment)
		前置条件	Comment为用户发布评论信息

	后置条件	数据库中保存comment信息，所有数据库同时更新
	不变量	无
需求接口		
PublishComment.addComment(Comment)		

ID： 47		构件： PlugAttackCheck, DistributedUpdate	
供给接口			
DistributedUpdate			
接口ID	功能描述	语法	void addIllegalIP(IP ip)
		前置条件	Ip为插件扫描攻击检测程序检测出的不合法ip地址
		后置条件	数据库中保存ip信息，所有数据库同时更新
		不变量	无
需求接口			
PlugAttackCheck.addIleagalIP(IP)			

ID： 48		构件： Monitor,NetworkService	
供给接口			
NetworkService			
接口ID	功能描述	语法	void ping(String serverID)
		前置条件	serverID为合法ID
		后置条件	将此ping请求加入到消息队列中准备发送
		不变量	无
需求接口			

```
Monitor.ping(ServerID)
```

ID: 49

构件: MonitorInterface,NetworkService

供给接口

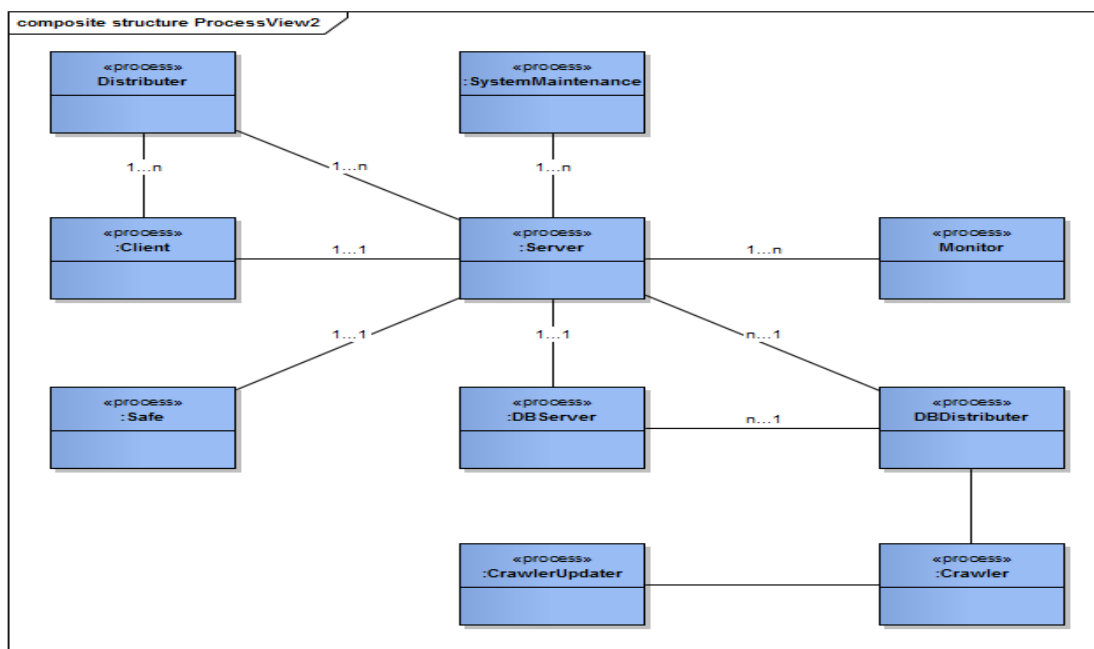
NetworkService

接口ID	功能描述	语法	void echo(String serverID)
		前置条件	此server正常工作, 且已经收到ping消息
		后置条件	将此消息放入消息队列中准备发送
		不变量	无

需求接口

```
MonitorInterface.echo()
```

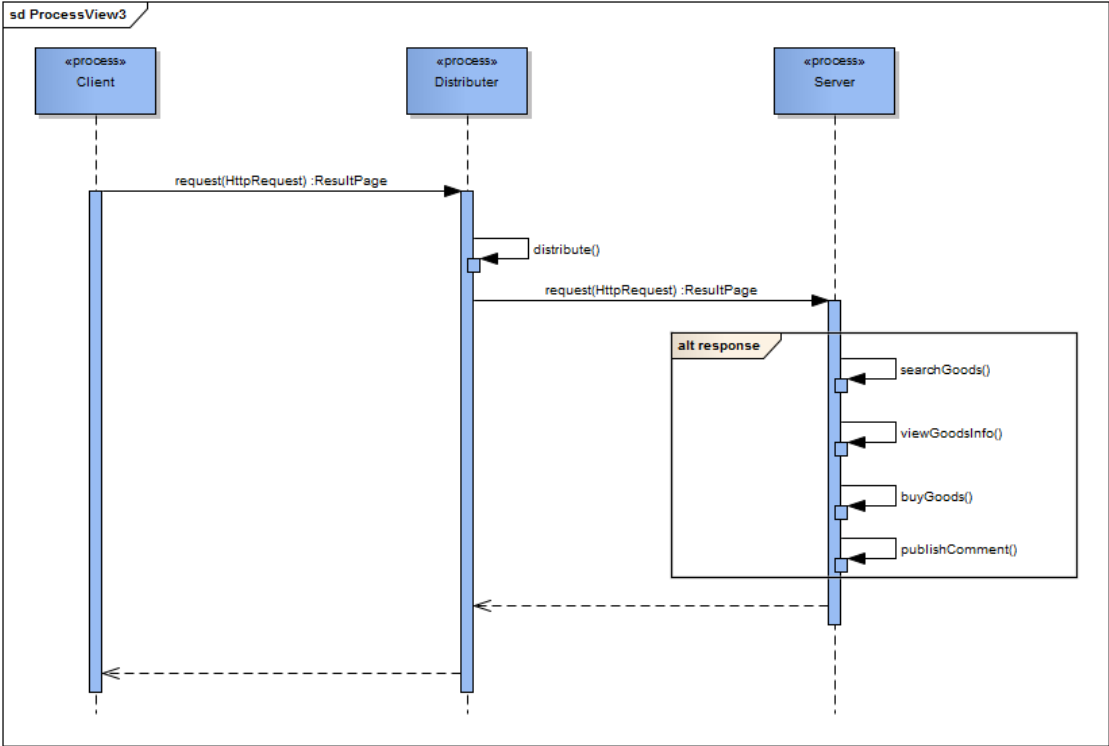
5.3 进程视图



逻辑视图与进程视图的映射关系描述

进程	部件/连接件
Client	View,BrowserCompatibler
Server	FrontController,SearchGoods,ViewGoodsInfo,BuyGoods,PublishComment,synProcissor,synDirectory,SortProcessor,sortRules,ShieldProcessor,ShieldDirectory,SearchProcessor,shareData
Safe	SafeGuard,AttackCheckAlgorithm,SIPCheck,PlugAttackCheck,shareData,Generation
System Maintenance	SystemMaintenance,CommentMaintenance,EventManager,Event,EventListener,IleagalUserCheck,ParticularCheck,ParticularDirectory,CheckComment,CheckAlgorithm,SensitiveCheck,SensitiveDirectory,Manage,Inform,Regist,Generation
DBServer	Data
Distributer	Distributer,DistributerRPC
DBDistributer	DataDistributer,DataBackup,DistributerRPC
Crawler Updater	UpdateScheduler,UpdateFraming
Crawler	Crawler,CrawlerConfig,CrawlerAlgorithm,UrlData,PageAnalysis,CrawlingData,StorageProcessor,CrawlingCompatibler,WebsiteACompatibler,PriceCompatibler,LanguageCompatibler,shareData
Monitor	Monitor,PingEcho

客户端、服务器通信进程描述

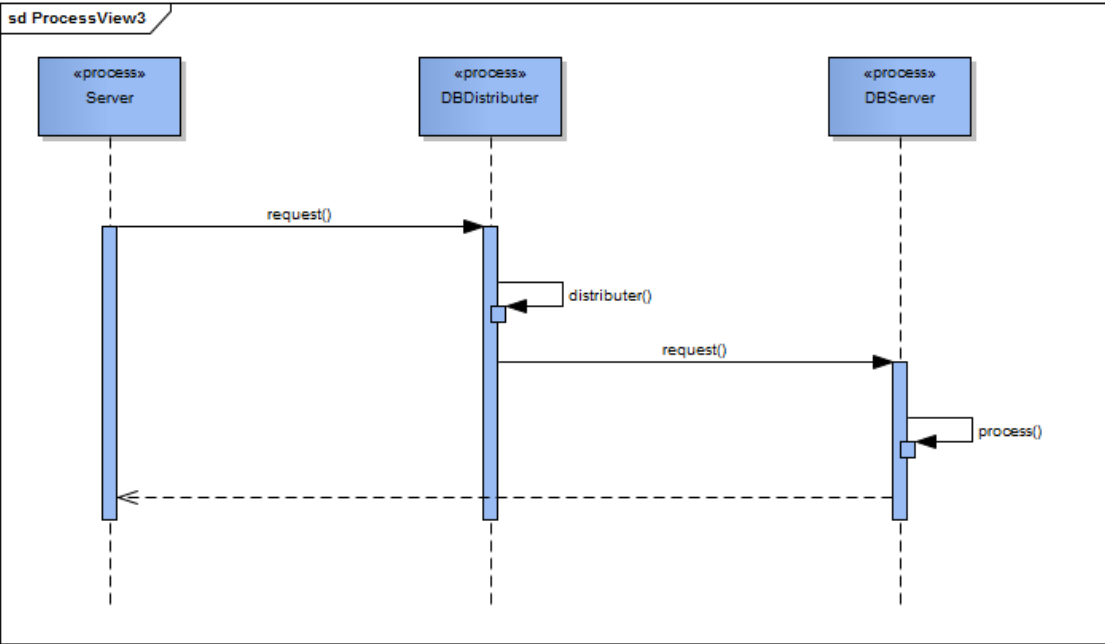


客户端、服务器通信进程接口描述

发送进程	Client
发送接口	request(HttpRequest)
通信方式	HTTP协议通信
通信描述	发送请求
接口进程/接口	Distributer.process()

发送进程	Distributer
发送接口	request(HttpRequest)
通信方式	HTTP协议通信
通信描述	分派请求
接口进程/接口	Server.process()

服务器、数据库通信进程描述

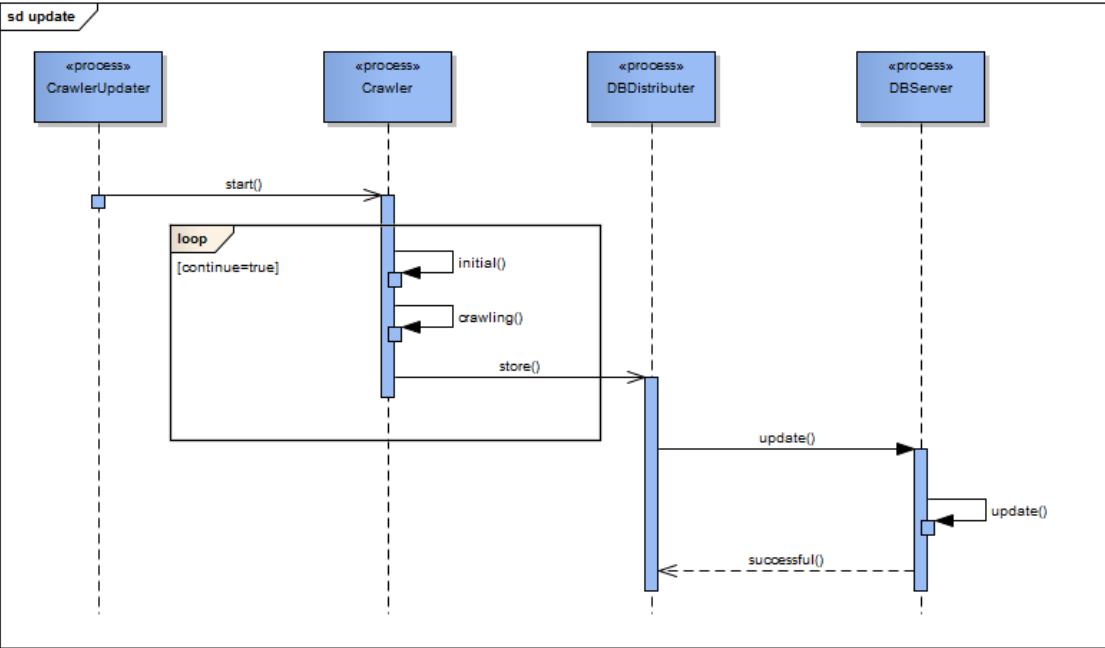


服务器、数据库通信进程接口描述

发送进程	Server
发送接口	request()
通信方式	Remote Procedure Call
通信描述	发送数据请求或更新数据库请求
接口进程/接口	DBDistributer.process()

发送进程	DBDistributer
发送接口	request()
通信方式	Cluster或Distribute通信
通信描述	分派请求
接口进程/接口	DBServer.process()

爬虫程序、数据库通信进程描述



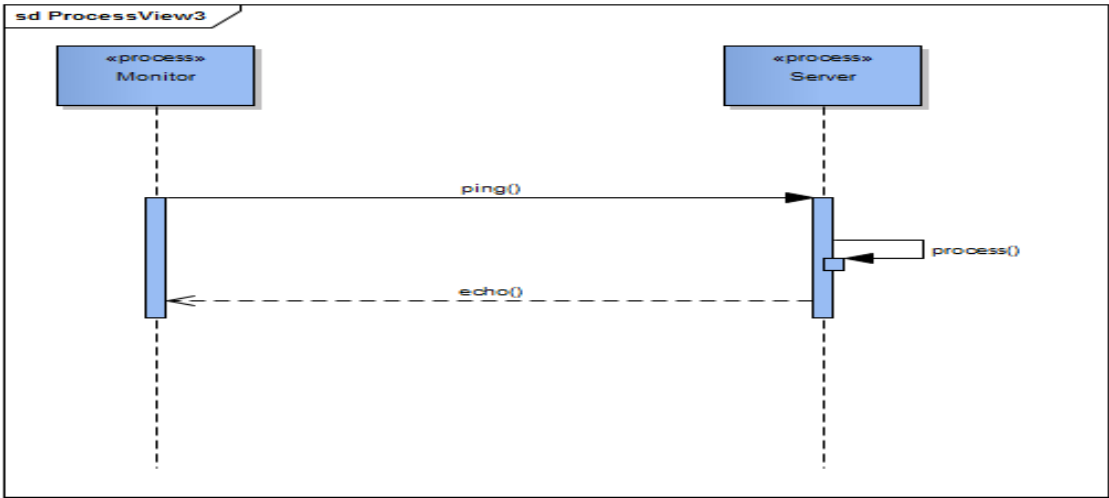
爬虫程序、数据库通信进程接口描述

发送进程	CrawlerUpdater
发送接口	Start
通信方式	Procedure Call
通信描述	建立爬虫程序
接口进程/接口	Crawler.initial()

发送进程	Crawler
发送接口	store()
通信方式	Remote Procedure Call
通信描述	请求向数据库保存数据
接口进程/接口	Distributer.process()

发送进程	DBDistributer
发送接口	update()
通信方式	Distribute通信
通信描述	更新数据库数据
接口进程/接口	DBServer.update()

服务器、监控的进程描述



服务器、监控的进程接口描述

发送进程	Monitor
发送接口	Ping()
通信方式	Remote Procedure Call
通信描述	发送ping消息监控server
接口进程/接口	Server.process()

发送进程	Server
发送接口	echo()
通信方式	Remote Procedure Call
通信描述	返回echo消息通知监控器
接口进程/接口	Monitor.process()

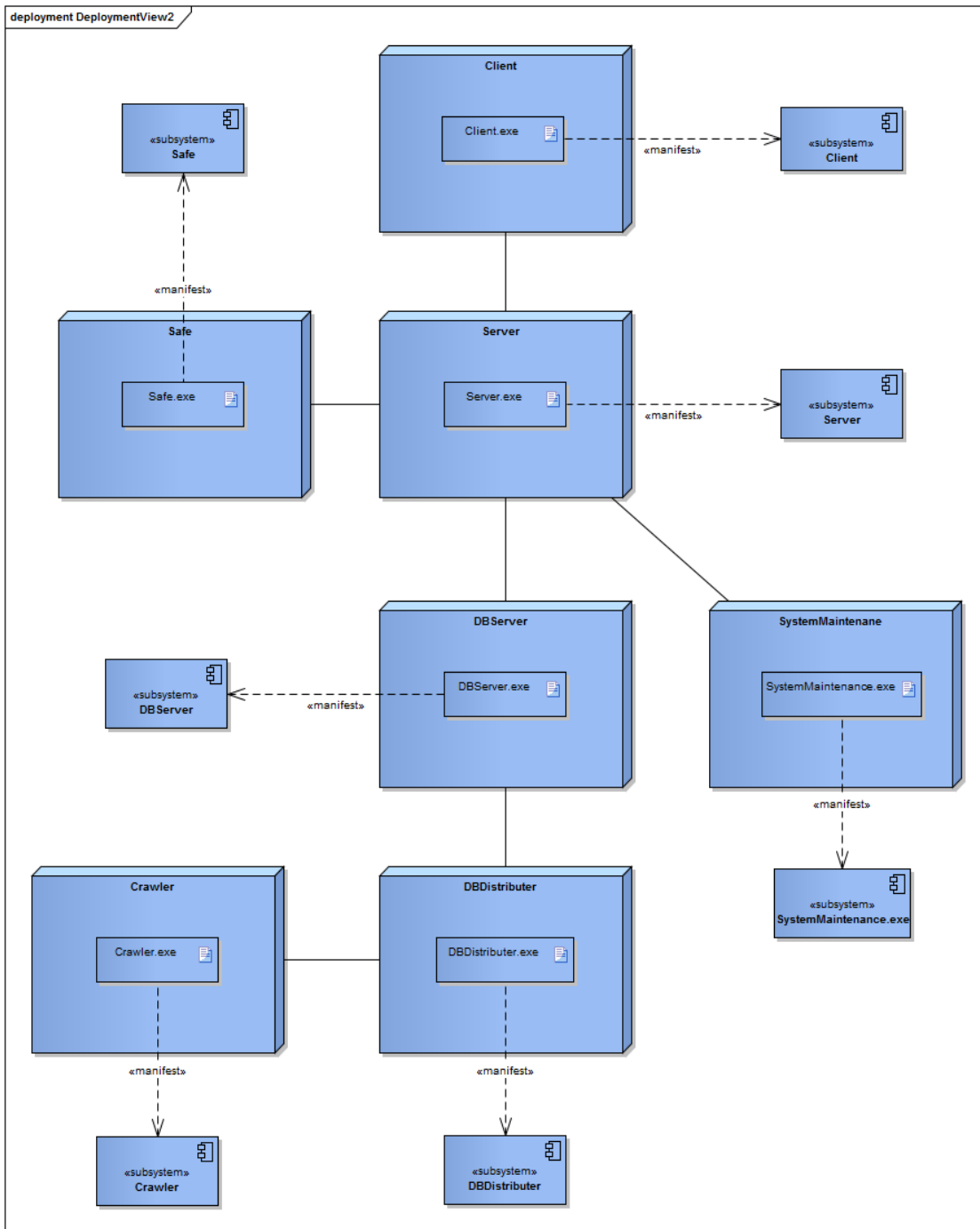
其他进程间通信接口描述

发送进程	Server
发送接口	commentMaintenance()
通信方式	Procedure Call
通信描述	请求进行评论维护
接口进程/接口	SystemMaintenance.process()

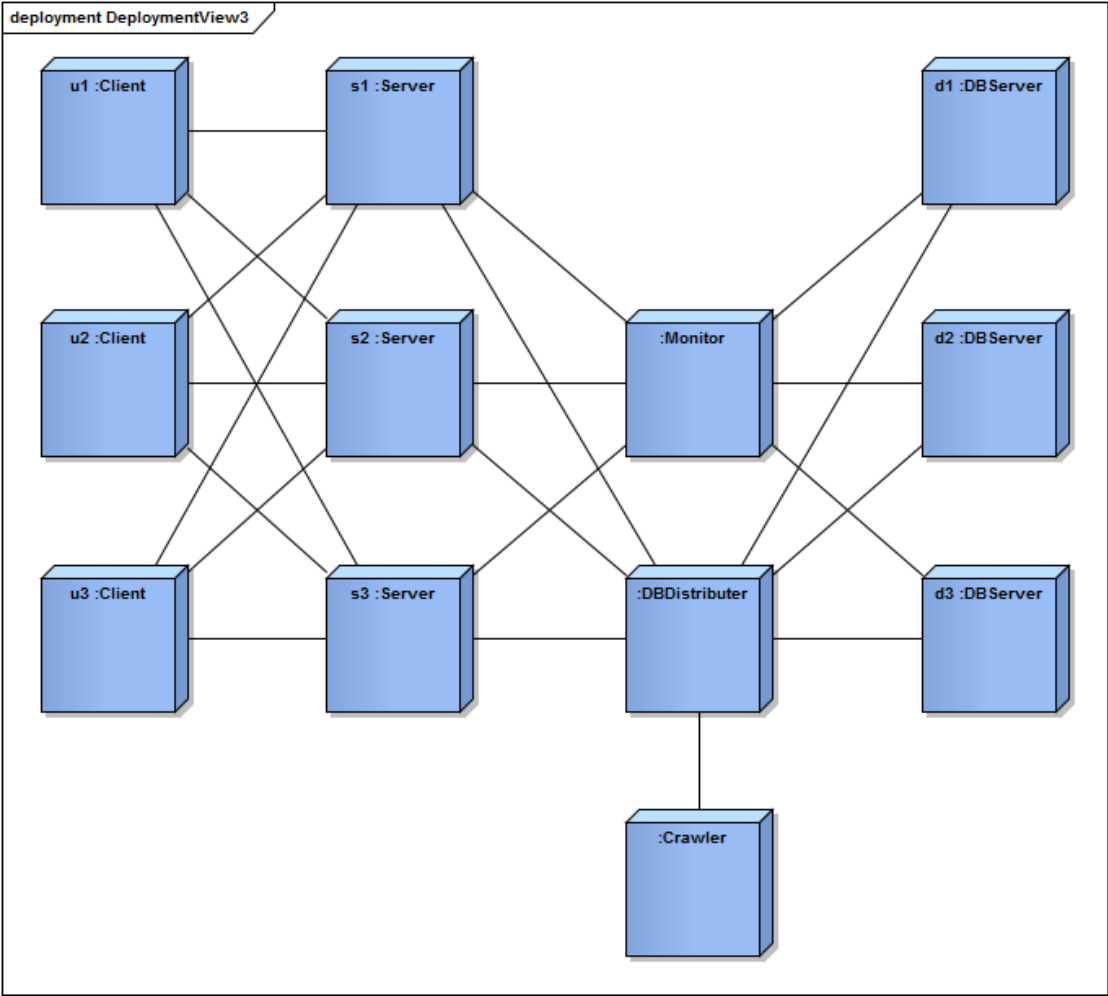
发送进程	Server
发送接口	attackCheck()
通信方式	Procedure Call
通信描述	请求进行入侵检测
接口进程/接口	Safe.check()

5.4 部署视图

比价系统的UML部署图描述



比价系统的基础设施分布描述



6.风格

6.1面向对象风格

使用编号		1
相关功能		高易用性搜索、搜索结果排序、屏蔽特定搜索结果、爬取商品信息、评论信息监测、水军检测、商家定制候选词汇
质量要求		允许定制同义词表、允许使用竞价规则定制排序策略、允许定制屏蔽商品列表、可自定义爬取来源网站、可以替换爬取策略，例如深度优先，广度优先、允许定制敏感词汇表、允许替换水军查找策略、允许商家定制特殊词汇表并通知商家
决策依据		决策1、2、3、4、5、8、12

使用编号	1
影响视图	

6.2 分层风格

使用编号	2
相关功能	商品搜索、查看商品信息、购买商品、发表评论、存储数据
质量要求	易用性搜索；评论信息敏感词检测；大存储量，最大能存储1千万个商品相关信息；反应时间最好在1s以内，1s-2s可接受；高峰期间能允许500万用户同时在线使用
决策依据	为了提高反应效率，负载均衡分配器将客户端请求发往不同服务器；为了提高容错性，数据分布存储；从而建立起从UIClient到Storage的五层结构
影响视图	

6.3 隐式调用风格

使用编号	3
相关功能	评论信息监测、水军检测、商家定制候选词汇
质量要求	允许定制敏感词汇表、允许替换水军查找策略、允许商家定制特殊词汇表并通知商家、增加新的评论检测规则代价小于0.25人月
决策依据	设计决策7
影响视图	