大纲（备注：所有图用visio画）

一、上一阶段工作

1.1 项目进展

自19年8月份“XXX”活动结束后，在修复活动中暴露的问题外，按照军科以及各地用户提出的需求，即时通信增加了较多功能，大致可以划分为：适配用户使用场景、增强应用稳定性、开发跨域功能、适配mmj、开发国产化版本5个方面。

**1.1.1 适配用户使用场景**

在XXX活动中以及之后的现场使用中，军科以及用户根据自身的使用模式提出了一些使用需求，即时通信增加如下功能：

a.发送文件前增加确认提示

b.统计一定时间内用户、要素之间的文字、文件传递数量

c.提供在web页面导出消息记录、群组消息的功能

d.单台电脑只允许运行一个即时通信进程，可在web端进行配置

e.增加讨论组的解散功能

**1.1.2 增强应用稳定性**

在2活动中即时通信暴露出一些稳定性问题，同时在19年12月的所内压力测试中暴露出即时通信服务器性能不足。针对这些问题，即时通信做了如下修改：

a.优化客户端传输大文件崩溃

b.优化文件传输变慢的问题

c.增加自动备份数据库以及自动删除大文件功能

d.开发服务器集群部署功能（已开发完成，尚未验证）

**1.1.3 开发跨域功能**

即时通信在2活动结束后的重点是开发跨域功能，采用中心服务器连接分域服务器的方式简化网络拓扑，用户的同步以及创建群组和群组消息通过中心服务器流转，共开发下述功能：

a.跨域组织架构同步，包含首次添加以及添加后的自动更新；

b.跨域单聊的发送信息，包含文字、文件、截图、表情、抖动、消息撤回等；

c.跨域群组的创建、管理以及群组发送信息；

d.状态消息跨域传递，包含用户在线状态、消息通知，群组内用户状态通知。

e.跨域音视频

**1.1.4 适配mmj**

自2活动结束后，即时通信共进行了3次mmj适配，分别是：

a.19年9月份在东部进行了消息加m的适配

b.20年6月份在南部进行了文件加m的适配

c.20年9月中巡在所内借用wd专项进行再次适配

至今为止即时通信和mmj只进行3次，时常不超过1个月适配，且只进行了功能的验证，性能如何没有经过详细测试。而ip加m方式完全未进行过适配。

**1.1.5 开发国产化版本**

即时通信目前正在同步开发国产化客户端版本。

1.2 当前软件架构

即时通信服务器由UpLoginServer、UpOrgServer、UpMessageServer、UpGroupServer、UpStatusServer、WebServer、antbiz多个服务器进程组成。下表1.2-0是各个服务器进程的功能说明。

|  |  |
| --- | --- |
| 名称 | 功能描述 |
| UpLoginServer | TCP连接挂载，跨域信令传输 |
| UpOrgServer | 用户登录，组织架构维护，用户角色权限相关信息维护 |
| UpStatusServer | 用户在线状态功能 |
| UpMessageServer | 单聊消息通知，存储，读取 |
| UpGroupServer | 群组操作及群组消息通知，存储，读取 |
| webServer | Web后台，系统设置，维护，配置，第三方API |
| antBiz | 组织架构更新，支持增量更新 |

表1.2-0

在单域部署的情况下，这些进程可以部署到独立一台物理机，也可以部署到多个物理机提供即时通信服务，如下图1.2-1小规模应用的情况下这些服务包括数据库可以运行在单个物理机上，目前实测单个高性能物理服务器支持日活5000人在线的规模。



图 1.2-0

在跨域的场景需求下我们将(图1.2-0)这样的一个服务器集合当成一个一个域，每个域具备独立的数据和完整的运行环境，域内客户机之间能实现完整的通信。跨需要跨域的时候，通过UpLoginServer实现服务器之间信令互通，从而实现跨域通信。如图1.2-1



图 1.2-1

**1.2.1 跨域互联架构**

**跨域互联可以实现多个独立的即时通信系统之间建立互联互通的能力。**

1.2.1.1 跨域群组实现

目前跨域互联架构为了解决多域互联问题，跨域群组及相关功能采用了中心互联的模式。由中心服务器来存储群组信息，群组的创建，删除，修改及分发群组的消息都通过中心服务器处理。



图1.2.1.1-0

图1.2.1-1是当前跨域中心互联模式的拓扑形式，客户机的跨域群消息通过本域的LoinServer服务器转发至中心服务器，相关的数据会被保存到中心服务器的数据库，并转发给指定域的客户机。

1.2.1.2 跨域单聊消息机制

客户发送单聊消息到本地LoginServer服务器，本地LoginServer根据目标用户的域ID直接将消息路由转发至目标域的LoginServer，目标域LoinServer将数据送至消息服务器处理。如图1.2.1-4 展示了A域将消息发送至B域的过程，这样设计的目的是为了在A域下线或者异常的时候，B域的客户端还能接收到。



图 1.2.1.2-0

1.2.1.3 跨域状态同步

跨域状态同步包括状态的拉取和广播，当一个用户登录上线的时候，不管是单域系统还是跨域系统，用户状态都会采取广播模式向所有人广播。

拉取在线状态：

当用户手动刷新或者登录完成的时候需要向本地服务器请求拉取一个完整的用户在线状态表。



图1.2.1.3-1

状态广播：

如图1.2.1-3-2当一个用户的状态变更的时候会向全部域所有客户发送广播通知。



图1.2.1.3-2

1.2.1.4跨域文件传输

跨域文件传输的实现原则是按需同步，这样某个域需要某个文件的时候才回去同步，不需要的话就不需要同步文件如图1.2.1.4-0。这样做的优点是节省带宽，和尽量减少同步。缺点是源服务器下线的情况，如果文件没有同步会无法收到文件。

跨域传输要点

* 跨域同步的时候只有第一个请求文件的用户可以同步接收，其他用户都需要等待同步
* 如果跨域同步服务器发生异常，crash等导致服务退出重启中断文件同步，服务器重启后如果继续同步应该能够断点续传
* 任何时候如果本地已经有某个文件应该从本地读取而不需要远程同步，比如A域和B域用户都同时上传了同一个文件的场景。



图1.2.1.4-0

1.2.1.2跨域服务管理

区域服务器安装好后，用户可以通过WEB后台设置好中心服务器地址，服务器会自动拉取中心服务器其他域的配置信息，然后将本地域的配置信息推送到中心和其他相关的域。用户可以根据需要选择需要互联的其他域。如图1.2.1-1



图1.2.1-2

1.2.1.3 跨域组织架构同步

所有区域与其他服务器首次建立连接，需手动拉取其他服务器的组织架构或者通过中心批量分发推送，当本地有其他区域的组织架构后，其他服务器增删人员，会自动同步到本区域

1.2.1.5 网络拓扑

跨域网络设计是两点之间UpLoginServer互联，但无法实现跨域路由。如下图，A、B、C三个域A B C都可以直连，但是无法做到A跨B转发信令到C。



图1.2.1.5-0

通过当前的互联模式，我们可以实现一下网络拓补模式。

1. 网状互联

如图1.2.1.5三个域互联的情况，我们需要一台中心服务器做跨域群组服务



1.2.1.5-1

1. 主从域模式

主域和中心服务器和所有域相连，从域之间不互联。



1.2.1.5-2

**1.2.2 加解m架构**

1.2.2.1 消息加m

当前跨域采取的是全信令加密模式，所有通过UpLoginServer跨域发送的信令都会被加密后发送，这里就包括消息信令。如图1.2.2.1-0



图1.2.2.1-0

1. 单聊消息加密

如图1.2.2.1-1展示了消息跨域加密的过程



图1.2.2.1-1

1. 群聊消息加密

如图1.2.2.1-2展示了当前群消息跨域加密的过程，因为消息内容被存储在了中心服务器，所以必须向中心服务器读取跨域群消息。



图1.2.2.1-2

1.2.2.3 文件加m

文件的加密只对文件内容的数据库进行加密，文件跨域同步时按需同步，当目标域需要某个文件的时候目标域向源域文件服务器发送同步请求启动同步过程。如图1.2.2.3-0



1.2.2.3-0

**1.2.3 程序改进**

1.2.3.1 分布式架构升级

对现有服务器进行分布式改造，以便即出现服务器性能不足的时候可以水平扩展。如图1.2.3.1-0是最终实现的分布式即时通信系统的架构图。



1.2.3.1-0

1. 分布式并行服务器设计

对当前的服务进行并行服务改造，主要改进前置UpLoginServer和后端的业务服务器UpStatuServer,UpGroupServer,UpMessageServer,UpOrgServer之间的通信关系. 下面以MessageServer为列介绍分布式计算单元的实现。改造的目标是实现一个可以多个UpLoginServer和多个UpMessageServer协同工作的可水平扩展，高可用的服务集群。

1. 剥离当前服务器中的内存缓存，移植到REDIS。

如图1.2.3.1-0将原来每个服务器的缓存移植到Redis保持缓存一致性，将内存和计算解耦剥离实现计算单元的并行。



图1.2.3.1-1

1. 实现信令分流和权重配置

改进后的UpLoginServer可以将信令根据权重分流到每个UpMessageServer处理。如图1.2.3.1-2 UpLoginServer将10条信令根据权重分配给了两台UpMessageServer.



图1.2.3.1-2

图1.2.3.1-3是客户机发送一条新消息信令流转的情况，信令可以在任何一台消息服务器处理，并准确将结果发送到客户端。



图1.2.3.1-3

1. 实现服务主动上下线动态扩展

服务上线：

通过WEB服务配置新上线的服务器，配置好后保存。新服务器的信息将推送给所有需要连接这台服务的其他服务。这些服务收到通知，自动和新服务器创建连接，在无需重启的情况下实现服务上线。

服务下线：

用户通过WEB服务器停止某台服务，相关服务应该收到通知并关闭网络连接，如果某台服务器异常关闭信令应该被分流到其他服务器处理，系统集群可以继续提供服务。

1. 分布式缓存设计

分布式缓存采用REDIS实现，关于RDIS的分布式部署实现可参考Rdis帮助文档。

1. 分布式存储设计

分布式存储采用Mysql，关于Mysql的分布式部署可以参考Mysql的帮助文档。

1. 分布式数据库设计

分布式存储采用MonogoDB，关于MonogoDB的分布式部署可以参考MonogoDB的帮助文档

1. 安装部署

为了保证安装部署方便，应该提供套件式样的部署能力。程序安装包应该具备完整的服务器组件，包括UpLoginServer , UpMessageServer, UpOrgServer, UpGroupServer,UpStatuServer,UpApiServer, WEB Server, REDIS, MongoDB, Mysql. 默认一键安装部署，可通过参数选择需要部署的服务，默认的包提供简单主从模式选项部署，如果需要更大规模分布式部署需运维处理。

如图1.2.3.1-4是一个简单的双机主从部署模式，两台服务器都部署了一套完整的IM程序通过一个前置的ngnix做TCP连接分流，通过安装选项将服务器部署成了主服务器和从服务器。更大的集群部署模式参考1.2.3.1-0。



图1.2.3.1-4

1.2.3.2 跨域代理服务

跨域代理主要是将跨域互联能力从先用域服务中剥离，以解决域之间耦合性过高的问题。跨域代理服务器应具备以下能力

1 服务自动注册，发现

2 信令跨域路由

3 服务器级联配置

4 黑名单，白名单功能

5 多服务器并行集群能力

6 中心管控能力，可远程开关代理

实现跨域代理之后域之间的互联和域即时通信服务隔离，可实现动态上下线可自动检测其他域是否在线，自动检测互联配置，路由配置。如果出现紧急情况需要阻断某个域进入域互联集群或者域紧急脱离跨域集群，可以直接断电关闭跨域代理。



1.2.3.2-0

1.2.3.3 消息内容JM

修改JM模式将全信令JM改成内容JM，并优化跨域群的转发模式降低中心服务器的JM压力。

1. 只对消息内容加密

1.2.3.3-0消息内容加密后，只对每条消息内容加密，极大降低了MMJ的压力



1.2.3.3-0

1. 跨域群主消息转发优化

如图1.2.3.3-1采用域扩散的方案，由消息上行服务器加密消息，消息到达中心服务器后将消息之间发送到需要改消息的其他域服务器。接收消息的域服务器收到消息后解包并对消息进行解密。这样中心服务器加无需加减密，域服务器各自解密一次即可。



1.2.3.3-1

1.2.3.4 文件服务器分布式存储

为了提升文件传输和存储能力，需对现有文件服务器增加负载均衡能力，以便在文件传输服务器性能不足的时候能动态增加文件传输服务器。简单的负载均衡如图1.2.3.4-1，每个文件服务器将自己的负责情况（正在进行的上传和下载个数）写入数据库，客户机通过HTTP接口访问WEB服务器查询并获得当前最空闲的文件服务，然后将文件传到该服务器。



图1.2.3.4-0

**1.2.4 当前问题及解决方案**

当前的架构在现场环境中遇到的一些问题，以及相关问题的原因。

1.2.4.1 本地备份文件架构

问题：导致登陆失败或者更新通信录失败

解决：优化客户端代码

1.2.4.2 通信录架构

问题：超过3000人通信录卡顿

解决：优化客户端代码

1.2.4.3 安装方式

问题：安装时可能出现丢失动态库

解决：优化安装脚本

1.2.4.4 跨域服务互相干扰

问提： 某个跨域服务不需要跨域了，其他域依然向该域发送信息

解决：参考1.2.3.2

1.2.4.5 服务器性能问题

问题：服务器压力大的时候CPU过高

解决：参考1.2.3.1

1.2.4.6 MMJ加密压力过大

问题：全协议加密MMJ压力过大

解决： 参考1.2.3.4

二、工作计划

2.1 活动中遇到的问题

**2.1.1 加解m瓶颈**

2.1.1.1 消息调用次数

2.2.1.2 文件加m大小

**2.1.2 网络拓扑限制**

**2.1.3 客户端性能瓶颈**

2.1.3.1 通信录瓶颈

跨域通信录加入后通信录结构变复杂，用户变多，会导致卡顿和崩溃

2.1.3.2 消息刷新瓶颈

聊天界面接收消息频率过大导致卡顿和崩溃

2.2 改进计划

**2.2.1 加解m改进**

51所提出了两种方式，ip和wdmmj，现在还不确定用那种，两种先准备着

2.2.1.1 使用wdmmj情况下只对正文加m

+ 架构

2.2.1.2 使用ipmmj情况下使用双网卡

+ 架构

**2.2.2 组织结构改进**

2.2.2.1 改进计划

针对当前同步组织结构时会同步对方所有用户，会拉取到一些本不需要他域可见的用户，导致用户组织结构臃肿，所以拟对用户进行角色配置，分为：

5级权限=域内通信能力

4级权限=5级+拥有跨域可见能力，可以跨域发送消息，不可以跨域发送文件

3级权限=4级+可以发送10M文件

2级权限=3级+可以创建互联群+可以发送50M文件

1级权限=2级+200M文件

2.2.2.2 改进方式

架构

**2.2.3 客户端改进**

2.2.2.1 本地缓存

2.2.2.1 通信录

2.2.2.3 安装方式

**2.2.4 消息路由**

使不直接相连的两个分域可以通过中继服务器消息跳转进行通信，只是构想，还需讨论。

**2.2.5 分层级架构**

中心服务器下挂一级分域，一级分域下挂二级分域，只是构想，还需讨论。