VM Optimization Plan

# Introduction

During the OS updates of 2012/03/27, it became clear that the locations of VMs on various production hosts needed to be optimized. For example, ticket1 is on ruckus, and ticket2 is on vm06, but although repo2 is on vm06, repo1 is not on ruckus but on vm04. This meant that while vm06 was being updated, neither vm04 nor ruckus could be updated without taking down both instances of the same service.

Our VMs are becoming quite disorganized. A little optimization could make a big difference.

# Goals

* VM hosts should either be production (i.e. every VM on it runs a production service) or nonproduction (i.e. no VM on it should run a production service).
* The instances of multi-instanced services should not be located on the same host, with rare exceptions (e.g. rsv-client).
* Minimize “update blackouts” by localizing dual-instanced services to pairs of servers.
* Minimize downtime of services that are internally necessary for updates.
* Prevent VM hosts from hosting too many VMs.
* Each production VM host should have a corresponding ITB VM host, on which are found the same services’ ITB instances.
* Minimize “special cases” that don’t follow the rules that most services follow.

# Summary of Production VM Hosts

The following is a list of currently active production VM hosts (the current state of affairs as of 2012/03/29).

## ruckus (vm-bl-01)

* Dell PowerEdge 2950
* Out of warranty since 2010/05/10
* Current host for display1, jira, lvs1, rsv-client1, rsv-client3, ticket1, tx1, www
* No well-defined associated ITB VM host

## ginrummy (vm-bl-02)

* Dell PowerEdge 2970
* Warranty ends 2013/05/27
* Current host for blogs, camon, ds-bl-02, gratiaweb1, jira-osg, lvs2, monitor
* No well-defined associated ITB VM host

## huey (vm-in-01)

* Dell PowerEdge 2950
* Out of warranty since 2010/05/10
* Current host for ds-in-01, jump2, lvs3
* No well-defined associated ITB VM host

## jazmine (devvm-bl-02)

* Dell PowerEdge 2950
* Out of warranty since 2010/05/10
* Current host for twiki
* No well-defined associated ITB VM host

## vm04 (dorothy)

* Dell PowerEdge R415
* Warranty ends 2014/06/03
* Current host for confluence, data1, myosg1, repo1, rsvprocess1, software1
* Has associated ITB VM host devm04 (lando) hosting confluence-itb, oim-itb, repo-itb, twiki-docteam

## vm05 (theitis)

* Dell PowerEdge R415
* Warranty ends 2014/06/03
* Current host for puppet-test, voms
* Has associated ITB VM host devm05 (kfcflu) hosting blogs-itb, ce, jira-osg-itb, rsv-dev, voms-itb

## vm06 (flagee)

* Dell PowerEdge R415
* Warranty ends 2014/06/03
* Current host for comet1, data2, event1, myosg2, oim, repo2, rsv-client2, rsvprocess2, software2, ticket2
* Has associated ITB VM host devm06 (ribbon) hosting comet-itb, event-itb, glidein-int, wn1, wn2

# Summary of Other ITB Hosts

These are ITB VM hosts that are not associated with a particular production host.

## grandad (devvm-bl-01)

* Dell PowerEdge 2950
* Out of warranty since 2010/05/10
* Current host for pakiti, rsv-client-itb, rsvprocess-itb, rsv-verify, software-itb, ticket-itb, and twiki-itb

## goodlove (devvm-bl-03)

* Dell PowerEdge 2970
* Warranty ends 2013/05/27
* Current host for data-itb, display-itb, gratiaweb-itb, jump, myosg-itb, rsv-itb, tx-itb, yum-internal

## freeman (vm-in-02)

* Dell PowerEdge 2970
* Warranty ends 2013/05/27
* Current host for adeximo, cpipes, echism, kagross, lvs4, rquick, soichi

## cindy (devvm-in-01)

* Dell PowerEdge 2970
* Warranty ends 2013/05/27
* Currently serving as glidein-itb, which will be moving

# Summary of Services

This is the current (as of 2012/03/29) state of the services that are on VMs (services on physical hosts are omitted).

I note that there are several services with a “1” that consist of only one instance. In cases where there is only ever going to be one instance, I would suggest that we remove the “1” and the DNS alias, if any, as the “1” only confuses matters.

## blogs

* SLA status: Production (no formal SLA)
* Instances: blogs on ginrummy
* ITB instance: blogs-itb on devm05

## camon (Doug Olsen’s CA monitor)

* SLA status: Production (unmaintained)
* Instances: camon on ginrummy
* No ITB instance

## ce

* SLA status: ITB
* Instances: ce on devm05
* No ITB instance

## comet

* SLA status: Production (low-priority)
* Instances: comet1 (there will be more eventually) on vm06
* ITB instances: comet-itb on devm06

## confluence

* SLA status: production (future)
* Instances: confluence on vm04
* ITB instance: confluence-itb on devm04

## data

* SLA status: production
* Instances: data1 on vm04, data2 on vm06
* ITB instance: data-itb on goodlove

## display

* SLA status: production
* Instances: display1 (there will be more) on ruckus
* ITB instance: display-itb on goodlove

## ds

* SLA status: production (internal)
* Instances: ds-bl-02 on ginrummy, ds-in-01 on huey
* No ITB instance

## event

* SLA status: production (low priority as this is a prototype)
* Instances: event1 (there will be more) on vm06
* ITB instance: event-itb on devm06

## gratiaweb

* SLA status: Production
* Instances: gratiaweb1 (unlikely that there will be more) on ginrummy
* ITB instance: gratiaweb-itb on goodlove

## jira

* SLA status: production (deprecated)
* Instances: jira on ruckus
* No ITB instance

## jira-osg

* SLA status: production
* Instances: jira-osg on ginrummy
* ITB instance: jira-osg-itb on devm05

## jump

* SLA status: production (internal)
* Instances: jump on goodlove, jump2 on huey
* No ITB instance

## lvs

* SLA status: production (lvs1/2), but lvs3/4 are unnecessary
* Instances: lvs1 on ruckus, lvs2 on ginrummy, lvs3 on huey, lvs4 on freeman
* No ITB instance

## monitor

* SLA status: production (internal)
* Instances: monitor on ginrummy
* No ITB instance

## myosg

* SLA status: production
* Instances: myosg1 on vm04, myosg2 on vm06
* ITB instance: myosg-itb on goodlove

## oim

* SLA status: production
* Instances: oim on vm06
* ITB instance: oim-itb on devm04

## pakiti

* SLA status: unnecessary (turn it off but keep it around)
* Instances: pakiti on grandad
* No ITB instance

## repo

* SLA status: production
* Instances: repo1 on vm04, repo2 on vm06
* ITB instance: repo-itb on devm04

## rsv-client

* SLA status: production
* Instances: rsv-client1 on ruckus, rsv-client2 on vm06, rsv-client3 on ruckus
* ITB instance: rsv-client-itb on grandad

## rsv-verify

* SLA status: unnecessary (turn off but keep around)
* Instances: rsv-verify on grandad
* No ITB instance

## rsvprocess

* SLA status: production
* Instances: rsvprocess1 on vm04, rsvprocess2 on vm06
* ITB instance: rsvprocess-itb on grandad

## software

* SLA status: production (passing into deprecation)
* Instances: software1 on vm04, software2 on vm06
* ITB instance: software-itb on grandad

## staff VMs

* SLA status: ITB
* Instances: adeximo, cpipes, echism, kagross, rquick, soichi
* No ITB instance

## ticket

* SLA status: production
* Instances: ticket1 on ruckus, ticket2 on vm06
* ITB instance: ticket-itb on grandad

## twiki

* SLA status: production
* Instances: twiki on jazmine
* ITB instances: twiki-itb on grandad, twiki-docteam on devm04

## tx

* SLA status: production
* Instances: tx1 (there may eventually be more) on ruckus
* ITB instance: tx-itb on goodlove

## voms

* SLA status: production
* Instances: voms on vm05
* ITB instance: voms-itb on devm05

## wn

* SLA status: ITB
* Instances: wn1 on devm06, wn2 on devm06
* No ITB instance

## www

* SLA status: production (deprecated)
* Instances: www on ruckus
* No ITB instance

# VM Host Pairs

I believe that each production VM host should have an ITB VM host that is similar to it in terms of hardware and OS, and that VM host should contain the ITB instances of services on the production host. We have this system, sort of, with vm04/5/6 and devm04/5/6, although the VMs could be better organized on these. However, some older hosts don’t have pairs, but could. My proposals:

## ruckus and grandad

They are of the same generation and have similar hardware.

## ginrummy and goodlove

These were purchased originally to be a production and ITB VM host.

# The Plan

Finally, the plan. Note that not everything in this plan will be implemented at the same time; this is just the final outcome.

## Create an Update-Critical VM Host

There are some services that are necessary for updates, and it is crippling when these services are down for OS updates. This chicken-egg problem can’t be solved 100%, but it can nearly be solved if we set aside a VM host (say, cindy, once glidein-itb is safely off it and on the server currently known as sonofsam) that can and will be updated out of band. As long as **no production services depend on anything on this server**, it is perfectly within our rights to update its OS and that of its VMs anytime that makes sense to us. ITB OS updates occur on the third Tuesday of each month; suppose we update this host on the Monday before that. There will then be no need for any of those services to be down during the regular ITB or production OS updates.

**Update-critical services include**: jump, monitor, yum-internal, yum-internal-6

## Shut Down Unnecessary VMs

Some VMs, as has been discussed, are no longer needed – and others might be added to the list.

**VMs to remove**: jump2, lvs3, lvs4, pakiti, rsv-verify

**VMs on their way out, but not yet**: jira (not jira-osg), software1, software2, www

**VMs about which there is an open question**: camon

## VM Host Alias Changes

ruckus, which is already named vm-bl-01, should be aliased as vm01 instead. grandad should be aliased as devm01 instead of devm-bl-01.

ginrummy, which is aliased as vm-bl-02, should be aliased as vm02 instead. goodlove is aliased as devm-bl-03, but it should be aliased to devm02. jazmine, which is currently aliased as devm-bl-02, should just have that alias removed. It is not well suited to be a VM host, and we are not allowed to put any VMs on it other than twiki anyway.

We may want to remove the aliases from huey, freeman and cindy, as they don’t really make sense anyway.

## VM Moves for Optimization

LVS pairs must be on the same VLAN, and we have several dual-instanced services that could run under LVS and probably will, so multiple instances of the same service must all be at one site.

Production services that are accessed by the outside world should be located at IUB.

## Highlights of the Changes

### Services moved for parallelism:

* rsv-client2, ticket2: moved to hosts which are home to the similarly-numbered instances of services whose other instance are on the parallel host (e.g. ticket1 is on vm01, so ticket2 should be on vm02)
* comet1, event1: moved to hosts which are home to other similarly-numbered instances
* blogs-itb, comet-itb, data-itb, display-itb, event-itb,tx-itb, jira-osg-itb, oim-itb, rsv-client-itb, software-itb, ticket-itb: moved to ITB hosts whose associated production host has a production instance of the same service

### Update-critical services moved to cindy:

* jump, monitor, yum-internal
* yum-internal-6 will soon be created for updating RHEL6 systems

### ITB VMs whose production instance isn’t on a VM

* rsv-itb has moved to devm05, where rsv-dev was already.
* twiki-docteam has moved to devm01, where twiki-itb was already.

### Other

* camon, Doug Olsen’s service, has moved to huey, which has become a lower-priority production host. This is assuming we continue to support camon, which we might not.

## Finally, the New Layout …

### Key:

Normal type: Instance remains where it was.

**Bold type**: Denotes the new home of an instance that has moved.

(Gray type in parentheses): Denotes the former home of an instance that has been moved or removed.

(Orange type in parentheses): Denotes the soon-to-be-former home of an instance that will be removed in the future, but we can’t remove it yet.

**(Bold orange type in parentheses)**: The instance is moving to this location for now, but it will be removed in the future.

**Green bold type**: Denotes a new instance that will be created in the future.

## vm01 (ruckus)

* Host for display1, (jira), lvs1, rsv-client1, rsv-client3, ticket1, tx1, (www)
* Associated ITB VM host devm01 (grandad) hosting **display-itb**, (pakiti), (rsv-client-itb), (rsv-verify), rsvprocess-itb, (software-itb), (ticket-itb), **twiki-docteam**, twiki-itb, **tx-itb**

## vm02 (ginrummy)

* Host for blogs, (camon), **display2**, ds-bl-02, gratiaweb1, jira-osg, lvs2, (monitor), **rsv-client2**, **ticket2**
* Associated ITB VM host devm02 (goodlove) hosting **blogs-itb**, (data-itb), (display-itb), gratiaweb-itb, (jump), **jira-osg-itb**, (myosg-itb), (rsv-itb), **rsv-client-itb**, **ticket-itb**, (tx-itb), (yum-internal)

## vm04 (dorothy)

* Host for **comet1**, confluence, data1, **event1**, myosg1, repo1, rsvprocess1, (software1)
* Has associated ITB VM host devm04 (lando) hosting **comet-itb**, confluence-itb, **data-itb**, **event-itb**, myosg-itb, (oim-itb), repo-itb, **(software-itb)**, (twiki-docteam)

## vm05 (theitis)

* Host for puppet-test, voms
* Has associated ITB VM host devm05 (kfcflu) hosting (blogs-itb), ce, (jira-osg-itb), rsv-dev, **rsv-itb**, voms-itb

## vm06 (flagee)

* Host for (comet1), **comet2**, data2, (event1), **event2**, myosg2, oim, repo2, (rsv-client2), rsvprocess2, (software2)
* Has associated ITB VM host devm06 (ribbon) hosting (comet-itb), (event-itb), glidein-int, **oim-itb,** wn1, wn2

## huey

* Host for **camon**, ds-in-01, (jump2), (lvs3)
* No well-defined associated ITB VM host

## freeman

* Host for adeximo, cpipes, echism, kagross, (lvs4), rquick, soichi
* No well-defined associated ITB VM host

## cindy

* Moves to IUB
* Host for **jump**, **monitor**, **yum-internal**, **yum-internal-6**
* Updated on the Monday before each ITB-OS-update Tuesday

# Comparison with Goals

* There are no longer ITB services running on production hosts or vice versa.
* The only cases where multiple instances of the same service are running on the same host are rsv-client and wn, which I believe to be the only exceptions. (This was already the case before the plan, however.)
* In order to avoid having both instances of a dual-instanced service down at the same time, I must merely avoid updating vm01 and vm02 at the same time, and avoid updating vm04 and vm06 at the same time.
* Updating cindy and its VMs on the Monday before we update ITB hosts will ensure that jump, monitor, and yum-internal are never down during regular updates.
* No VM host has more than 8 services on it.
* ITB instances are all on hosts that are partnered with a production host that runs the production instance of that service, where applicable.
* Unfortunately this does nothing to reduce the number of “special cases” (camon, twiki, etc.)

# Implementation

The plan could be implemented in stages, but I’m leaving that planning until after others have had a chance to look at the plan, as there’s no point in discussing the minutiae of a plan that could easily change.

There are 20 VM moves here, including one that is changing from KVM to VMware, which is possible but which I’ve never done before (I’ve gone the other direction a few times, though).

There’s still no vm03 or devm03; perhaps there will be eventually.