Discovery
OpenStack Fault Management Framework

Nemat Bidokhti

Contributors:

- Farhad Sunavala (Huawei)

Agenda

- **Introduction**
- Fault Synopsis
- Deep-dive: Analysis Workflow
- Future Deliverables
- Asks

Motivation

- Many companies now depend on OpenStack to operate their infrastructure. As a distributed project, consisting of many contributors from around the world, standards for product quality and reliability may be inconsistent.
- Introduce uniformity to OpenStack design and deterministic operation
- Reduce MTTR
- Build a fault management knowledge database

Vision & Mission

- Facilitate development and deployment of OpenStack as a business-critical application with quality and resiliency that meets or exceeds commercial product standards.

- Contribute to the education and training of community members, with respect to fault detection and remediation, along the following themes:
  - ✔ Build a better community
  - ✔ Build a better operator experience
  - ✔ Build a better product

Current Work

Phase 1: Identification

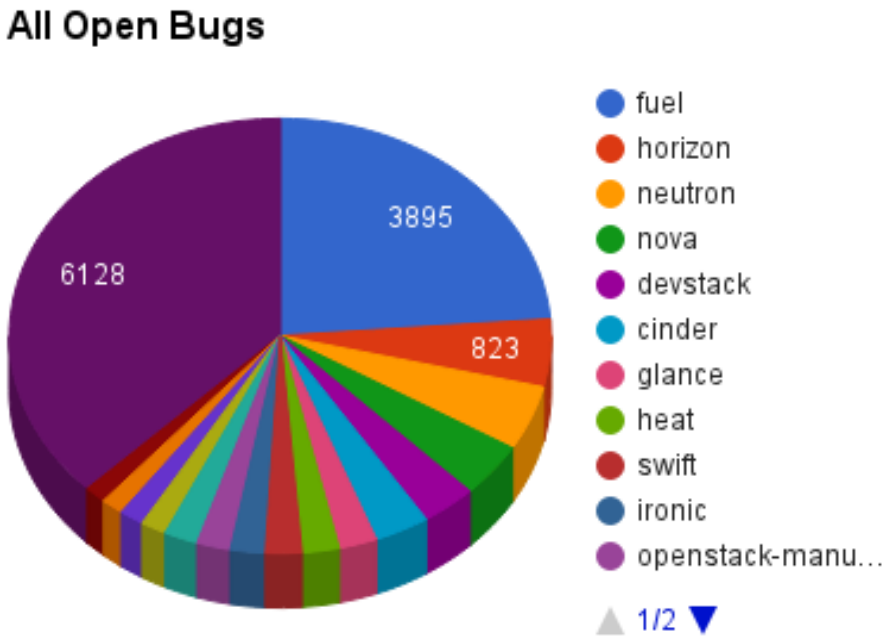
- Collect as many fault patterns as possible
  1. Survey the community
     - Manual interviews with a focus on OpenStack operators
  2. Data-mining exercise
     - Pull data programmatically from launchpad and stackoverflow

- Categorize reported faults with two distinct goals
  1. Construct a template for the purposes of facilitating collection and analysis of faults
  2. Produce a report to the community to share findings

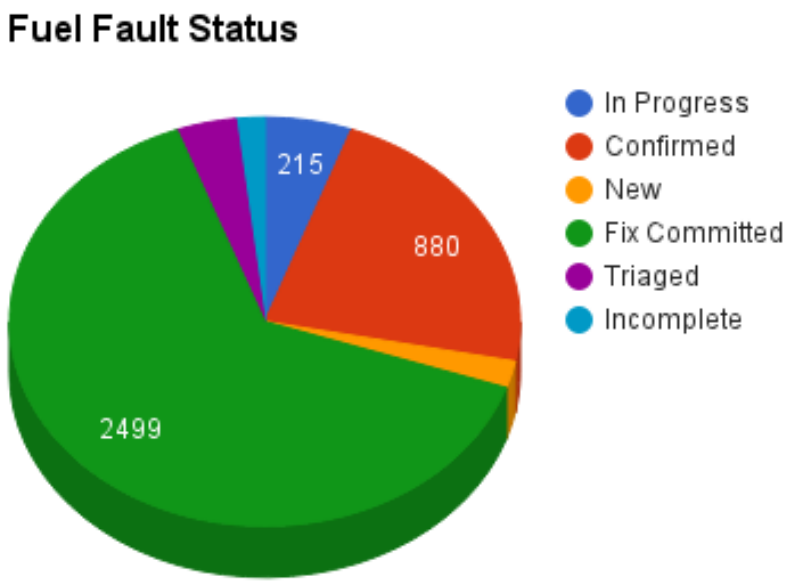
Agenda

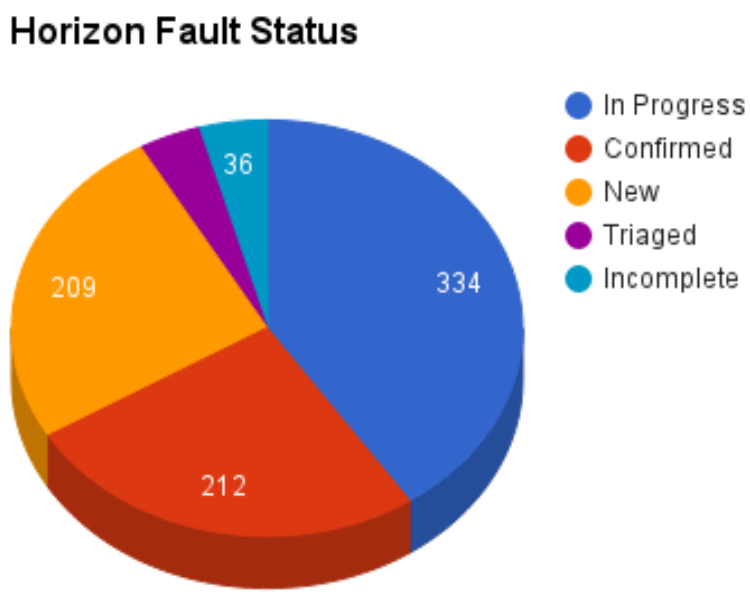
- Introduction
- **Fault Synopsis**
- Deep-dive: Analysis Workflow
- Future Deliverables
- Asks

OpenStack bug summary

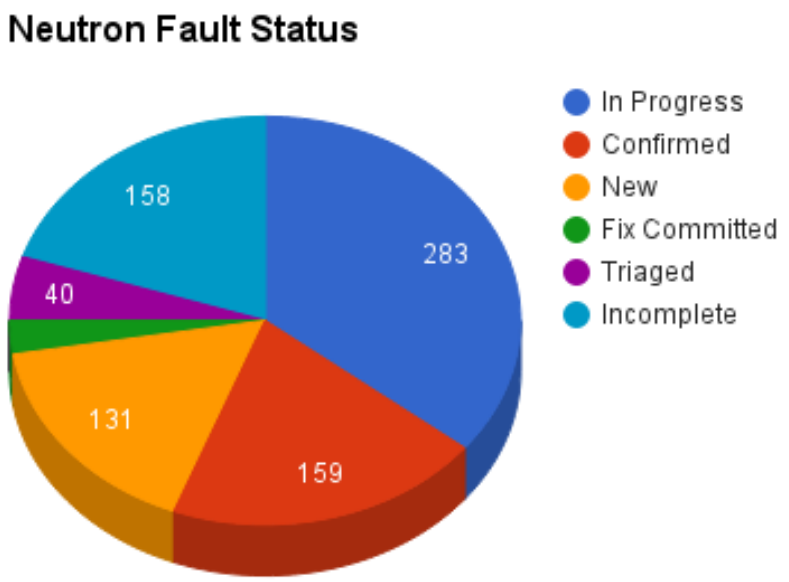


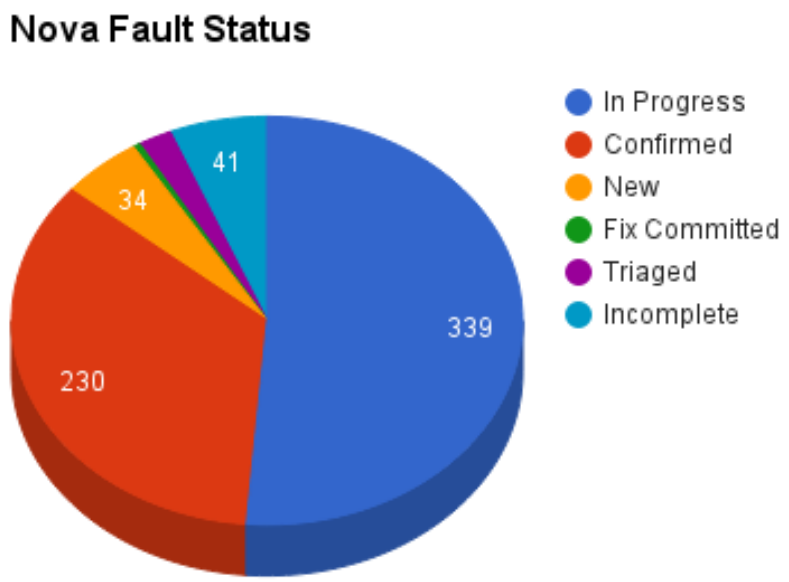
Top 2 Projects



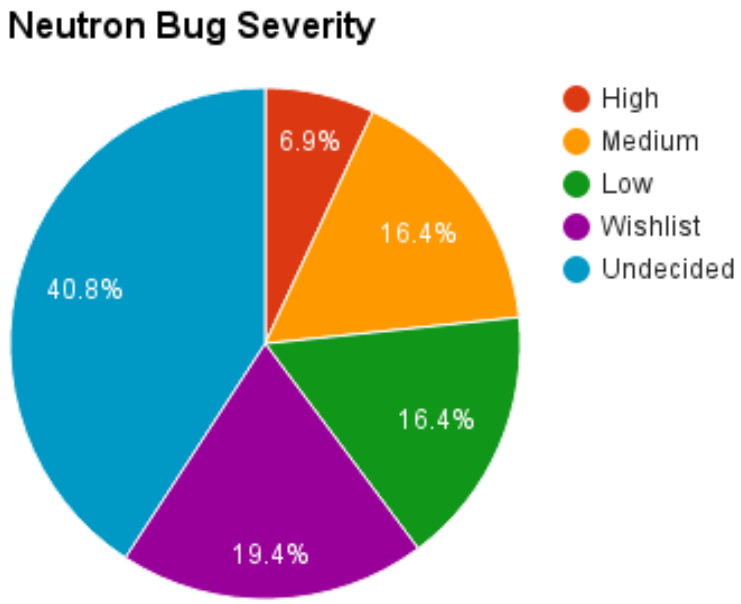


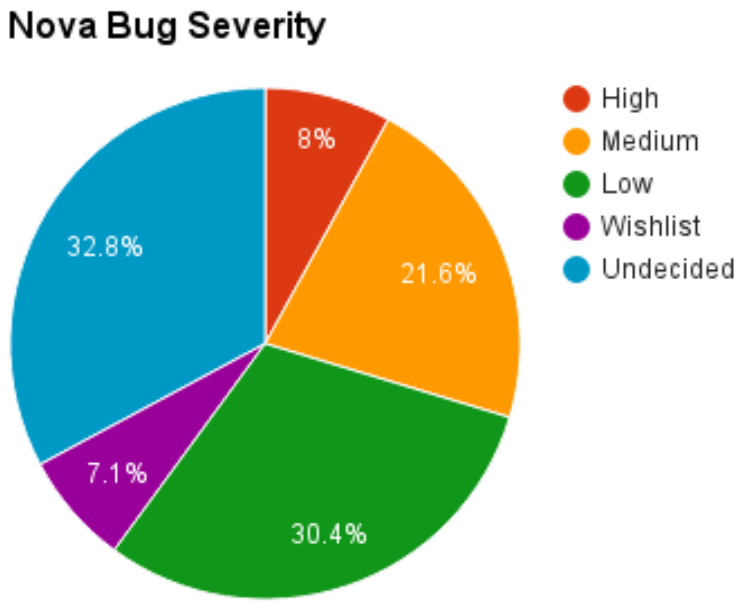
Next 2 Projects





Severity Distribution





Agenda

- Introduction
- Fault Synopsis
- **Deep-dive: Analysis Workflow**
- Future Deliverables
- Asks

Fault Genes Worksheet

Currently it is in Excel format and resides in Google Doc

- Req
- Bug
- Version
- Project
- Component
- Fault Class
- Fault Types
- Fault Sub-Type/Root Cause
- Fault Description
- Fault Symptom
- Severity
- Priority
- Status
- Mitigation
- Log
- Repro
- Submitter
- Assignee
- Created
- Deployment
- Fault Insertion Method

Fault Genes Workflow (Data from Launchpad)

req	project	fault_description	severity
all sfc flows gone after restart ovs agent	networking-sfc	In my setup, i create a port-chain and in br-int, there are some flows about sfc, after i restart the neutron-openvswitch-agent, all the sfc flows gone.	High
HA router should failover if GW address is not reachable	neutron	If a HA router has an external interface defined then it should use a custom keep lived health check to monitor the default gateway. If it isn't reachable after X attempts, lower its own priority and failover.	High
Problem in l3-agent tenant-network interface would cause split-brain in HA router	neutron	Assuming l3-agents have 1 NIC (ie eth0) assigned to tenant-network (tunnel) traffic and another (ie eth1) assigned to external network,. Disconnecting eth0 would prevent keeplived reports and trigger one of the slaves to become master. However, since the error is outside the router namespace, the original master is unaware of that and would not trigger "fault" state. Instead it will continue to receive traffic on the, yet active, external network interface - eth1.	High
Ipset race condition	neutron	Hello, We have been using ipsets in neutron since juno. We have upgraded our install to kilo a month or so and we have experienced 3 issues with ipsets. The issues are as follows: 1.) Iptables attempts to apply rules for an ipset that was not added 2.) iptables attempt to apply rules for an ipset that was removed, but still refrenced in the iptables config 3.) ipset churns trying to remove an ipset that has already been removed. For issue one and two I am unable to get the logs for these issues because neutron was dumping the full iptables-restore entries to log once every second for a few hours and eventually filled up the disk and we removed the file to get things working again.	High

Fault Genes Workflow (Brainstorming Phase)

req	project	fault_class	fault_type	root_cause	fault_description	severity	mitigation
all sfc flows gone after restart ovs agent	networking-sfc	Service Chaining Feature Not Working	Agent Issue	Timing Issue	In my setup, i create a port-chain and in br-int, there are some flows about sfc, after i restart the neutron-openvswitch-agent, all the sfc flows gone.	High	
HA router should failover if GW address is not reachable	neutron	Failover Issue	Gateway Access Not Reachable		If a HA router has an external interface defined then it should use a custom keep lived health check to monitor the default gateway. If it isn't reachable after X attempts, lower its own priority and failover.	High	Use Custom Keep alive Health Check
Problem in l3-agent tenant-network interface would cause split-brain in HA router	neutron	Failover Issue	Design Issue of OpenStack Network Node		Assuming l3-agents have 1 NIC (ie eth0) assigned to tenant-network (tunnel) traffic and another (ie eth1) assigned to external network,. Disconnecting eth0 would prevent keeplived reports and trigger one of the slaves to become master. However, since the error is outside the router namespace, the original master is unaware of that and would not trigger "fault" state. Instead it will continue to receive traffic on the, yet active, external network interface - eth1.	High	Out of Band Health Checking Mechanism
Ipset race condition	neutron	Security	Security Group Mis-Configuration	The Security Group	Hello, We have been using ipsets in neutron since juno. We have upgraded our install to kilo a month or so and we have experienced 3 issues with ipsets. The issues are as follows: 1.) Iptables attempts to apply rules for an ipset that was not added 2.) iptables attempt to apply rules for an ipset that was removed, but still refrenced in the iptables config 3.) ipset churns trying to remove an ipset that has already been removed. For issue one and two I am unable to get the logs for these issues because neutron was dumping the full iptables-restore entries to log once every second for a few hours and eventually filled up the disk and we removed the file to get things working again.	High	Change Security Rules/Programming the Security Group

Fault Genes Workflow (Brainstorming Phase)

req	project	fault_ class	fault_type	root_ cause	fault_description	severity	mitigation
neutron should not try to bind port on compute with hypervisor_type ironic	neutron	Networking Issue	Bare Metal Not Accessible Via Network		Neutron tries to bind port on compute where instance is launched. It doesn't make sense when hypervisor_type is ironic, since VM does not live on hypervisor in this case. Furthermore it leads to failed provisioning of bare metal nod e, when neutron is not configured on ironic compute node. Setup: node-1: controller node-2: ironic-compute without neutron neutron-server.log: http://paste.openstack.org/show/445388/	High	Fix in Code
Neutron subprojects cannot depend on alembic revisions from other neutron subprojects	neutron	Design Issue			If networking-foo depends on networking-bar, then a requirement may arise for an alembic revision in networking-foo to depend on an alembic revision in networking-bar. Currently this cannot be accommodated because each subproject has its own alembic environment. To solve this issue we need to switch to one alembic environment (neutron's) for all neutron subprojects.	High	
lbaas:after create 375 LB pool , the new lb -pool and vip get in error status	neutron	Scalability Issue	Message Timeout/Queue Overflow		1 create two-arm LB with 1client and 1 backend server on a tenant 2 repeat step1 to create 375 tenants 3 after step 2 , the LB network unstable	High	Improve Architecture Design
LBaas-LB performance just have 1 G low performance than LB bypass have 4G	neutron	Performanc e Issue			LB performance just have 1G low performance than LB bypass have 4G setup infor for LB bypass , client directly send traffic to server without LB , we have 4G performance and for LB, client send traffic with LB , we just have 1G traffic so LB is a bottleneck	High	

Fault Genes Workflow (Brainstorming Phase)

req	project	fault_ class	fault_type	root_ cause	fault_description	severity	mitigation
router:dhcp ports are open resolvers	neutron	Security	DNS Attack		When configuring an public IPv4 subnet with DHCP enabled inside Neutron (and attaching it to an Internet-connected router), the DNS recursive resolver service provided by dnsmasq inside the qdhcp network namespace will respond to DNS queries from the entire Internet. This is a huge problem from a security standpoint, as open resolvers are very likely to be abused for DDoS purposes. This does not only cause significant damage to third parties (i.e., the true destination of the DDoS attack and every network in between), but also on the local network or servers (due to saturation of all the available network bandwidth and/or the processing capacity of the node running the dnsmasq instance). Quoting from http://openresolverproject.org/:	High	Implement CERT Alerts Updates
Timeouts in update_device_list (too slow with large # of VIFs)	neutron	Scalability Issue	Message Timeout/Queue Overflow		In our environment, we have some large compute nodes with a large number of VIFs. When the update_device_list call happens on the agent start up: https://github.com/openstack/neutron/blob/master/neutron/plugins/ml2/drivers/openvswitch/agent/ovs_neutron_agent.py#L842 This takes a very long time as it seems to loop on each port at the server side, contact Nova and much more. The default rpc timeout of 60 seconds is not enough and it ends up failing on a server with around 120 VIFs. When raising the timeout to 120, it seems to work with no problems.	High	Implement CERT Alerts Updates
ipv6 neighbor advertisement storm	neutron	Networking Issue	Network Design Issue	Network Storm	stable liberty cloud with 20 network nodes, running OVS and supporting 1200 projects. Each project has on network with IPv4 and IPv6 subnets and one project router to attach to the external network. Network nodes are seeing 1000 IPv6 Neighbour Advertisements within 2.3 seconds.	High	Reduce L2 Broadcast Domain

Fault Genes Workflow (Neutron Failure Modes Samples)

Fault Class	Fault Types	Root Cause	Mitigation
VM is up not Accessible via network	network connectivity issues	Virtual interface in the VM admin down	Un-shut the virtual interface
		Virtual interface does not have IP address via DHCP	Depends on lower level root cause
		Virtual network does not have interface to the router	Add virtual network as one of the router interfaces
		vNIC port of VM not active (stuck in build)	Depends on lower level root cause
		Security group lock in traffic	Fix the security group to allow relevant traffic
		Configuration error	Isolate the problem to drivers/agent & look for possible errors in the respective config files (for a virtual overlay network the VNI range is not specific in the ML2 config, instead the operator has mistakenly configured the VLAN rang)
		SSH keys not correct	
Not Accessible via VNC	Web server not set up properly on the controller		Install web server correctly on the controller
	VNC config error		fix configuration
VM in Error State	Unable to Add Port to Bridge	Libvirtd in Apparmor is blocking	allow Libvirtd profile in Appamor
	No Valid Host Found/insufficient hypervisor resources	Compute nodes do not have sufficient resources	free up required compute storage and memory resources on compute node
	No Resource	Configuration issues	Change config setting
	Failed to Allocate Networks		
	authentication/permissions error	Configuration error such as port # or Password	Make sure end points are properly configured
Service Chaining Feature not working	Agent issue	Timing issue	
Failover issue	Gateway access not reachable		Use custom keep-alive health-check
	Design issue of OpenStack Network node		Out of band health checking mechanism
Security	Security Group Mis-configuration	The security group	Change security rules/Programming the security group
	DNS Attack		Implement CERT alerts updates
Networking issue	Bare metal not accessible via network		Fix in code
	Network design issue	Network storm	Reduce L2 broadcast domain
Design issue			
Scalability issues	Message timeout/queue overflow		Improve architecture design
Performance issue			

Agenda

- Introduction
- Fault Synopsis
- Deep-dive: Analysis Workflow
- **Future Deliverables**
- Asks

Future Deliverables

Phase 2: Education

- Periodic Report to the community
- Operator's Guide to Fault Detection, Reporting and Remediation
- Implementor's Guide to Fault Detection and Reporting
- Summit workshop(s) to present case studies, findings, technical work products

Future Deliverables

Phase 3: Productization (Customers)

- Automatic/simplified issue generation to facilitate operator-reported faults
- Standardized APIs for fault reporting by OpenStack software components
- Consistent /Deterministic fault behavior
- Health Manager framework to facilitate self-diagnosis of operational status
- Architectural and design patterns that focus on robustness

Agenda

- Introduction
- Fault Synopsis
- Deep-dive: Analysis Workflow
- Future Deliverables
- **Asks**

How to Get Involved?

- ✔ We can only learn and share when bugs are reported to the community
- ✔ There are a lot of bugs to analyze
- ✔ Need experts on various OpenStack components

- Contact us
  - Nematollah Bidokhti Nematollah.Bidokhti@huawei.com
  - Matt Greene, https://launchpad.net/~m-greene

- Join the working group
  - Weekly meetings on Monday

Thank you

www.huawei.com