

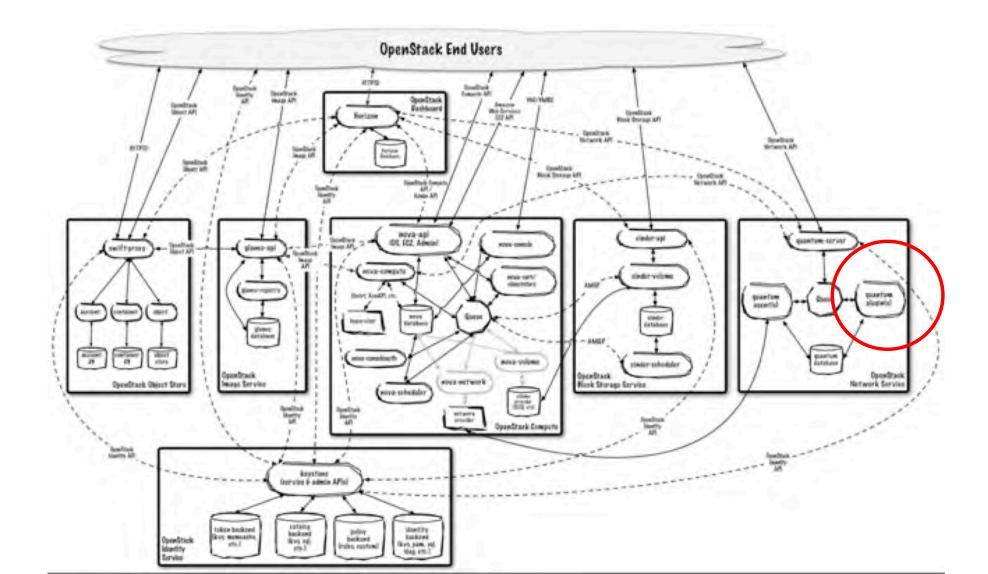
Debugging Openstack Problems Using A State Graph Approach

Yong Xiang, Hu Li, Sen Wang, Charley Peter Chen and Wei Xu

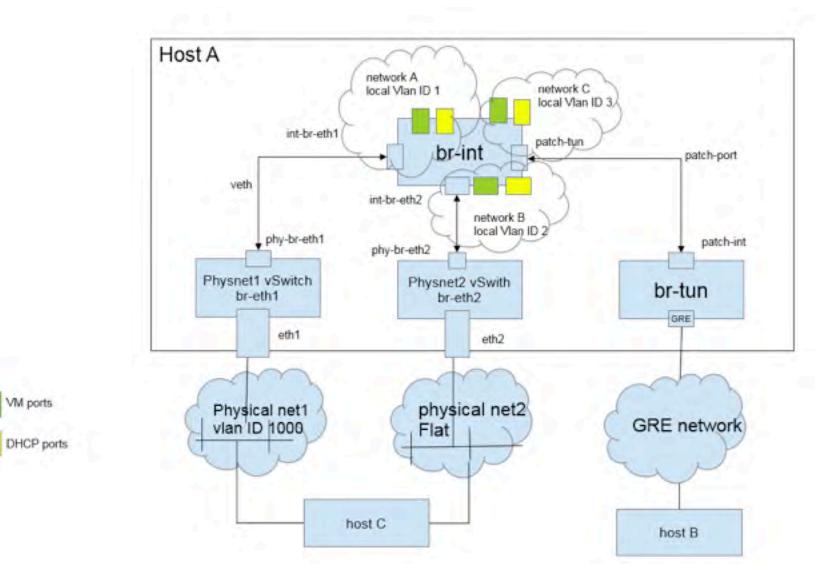
Institute for Interdisciplinary Information Sciences (IIIS),

Tsinghua University

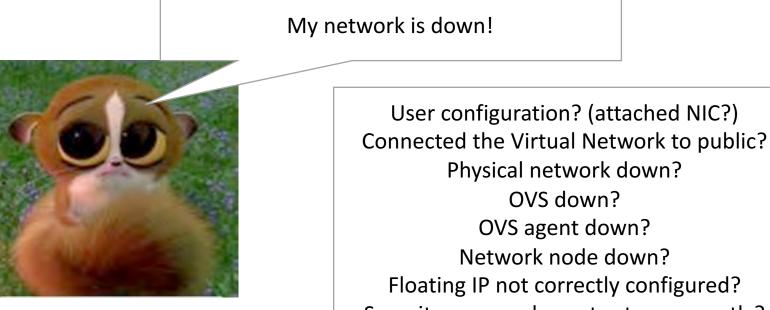
Modern systems are complicated



Modern systems are complicated (cont'd)



Trouble shooting for clients

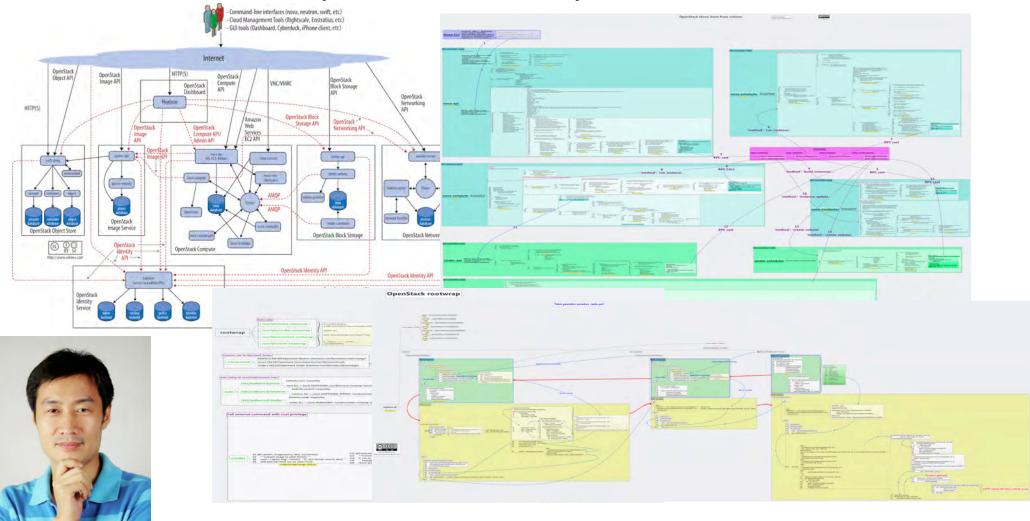


Security group rules not set up correctly?

••••



How many rules needs to know as a professional openstack operator?



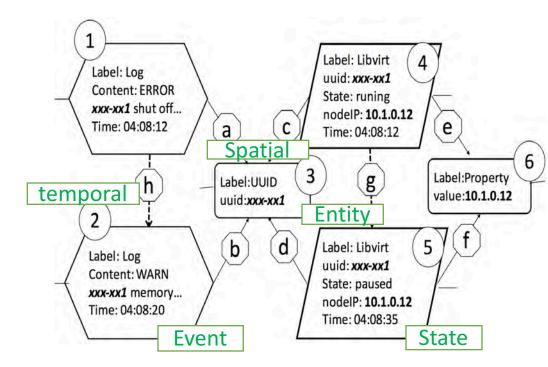
The operational knowledge does not transfer!

Good news for IT consulting business.



Key idea: automatically discover knowledge in systems using most basic rules

- We can capture the knowledge: System Operation State Graph (SOSG)
- Turn ad-hoc system state queries into a uniform graph traversal.
- Anomaly detection to find hidden problems.



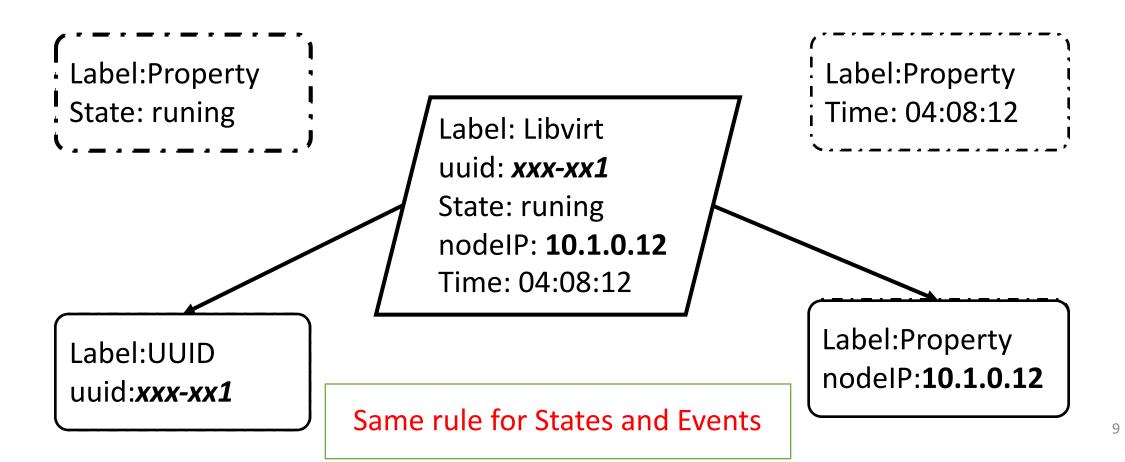
Data source used to construct the graph

• raw data sources, no semantic information

Туре	Data source	_	
DB	OpenStack databases updates (triggers)		
Libvirt	libvirt status Python API		
Ovs	OVS status ovsdb-client dump	States	
Cephimage	Ceph image list rbd info		
Cephfile	Ceph block file ls /ceph/dir/file		
Cephlog	log files from Ceph		
Log	logs from all OpenStack components	Events	

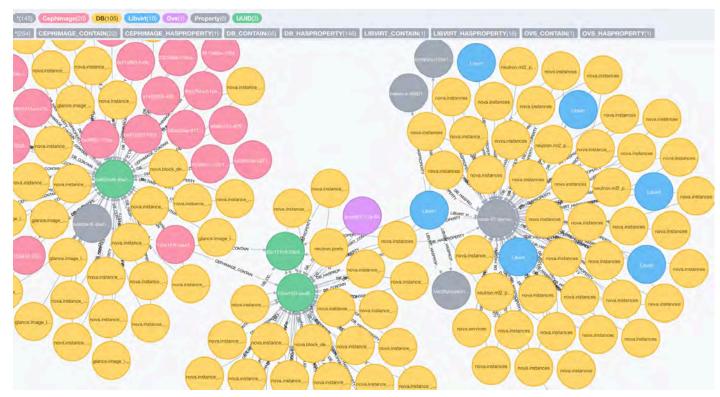
How to construct the graph?

• Simple rule: discovery entity based on syntactic IDs



Can be very a large graph!

- 3-day operation data, about 40 GB
- Graph size: 43.3 million vertices, 56.6 million edges

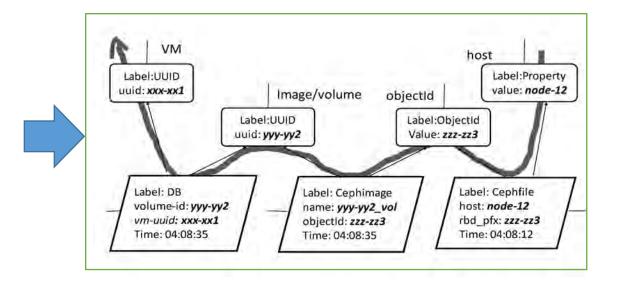


How can we use the graph?

- System query as graph traversal
 - Ad-hoc queries => uniform method
 - No need to memorize tons of system-specific commands
- Anomaly detection
 - Automatically find the hidden problems in system with millions of states

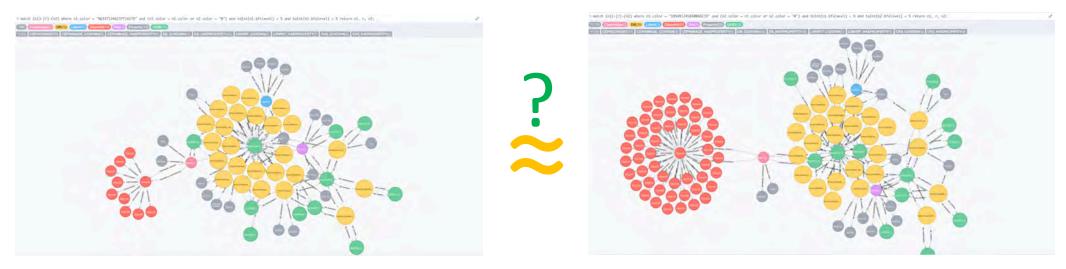
System query as graph traversal

- If physical server A encounters a hard disk failure, which VMs are affected? (Ceph as the Openstack storage backend)
 - 1. Which blocks are stored on the disk (Linux) *ls /var/lib/ceph/osd/...*
 - 2. Which ceph image the block belongs to (Ceph) rbd info -p compute(or volumes) rbd info -p compute(or volumes) <image> grep block-name-prefix filename
 - 3. Where the image is used (Openstack) nova show <server> nova volume-show <volume> cinder show <volume>



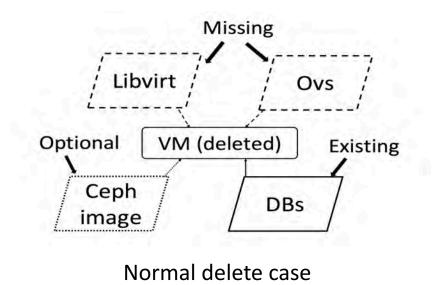
Anomaly detection: ideas

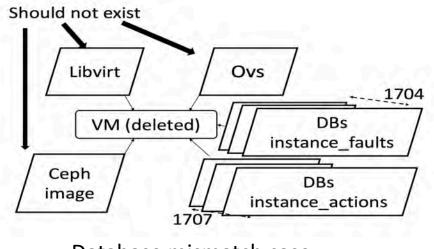
- Based on subgraph describing a single VM
 - find the subgraph that roots at the VM and also includes all its dependencies
- Distance-based anomaly detection
 - capture the structure information of the subgraphs in the distance metric



Anomaly case study: database record does not match physical states

- Subgraph: thousands of DB state vertices directly connect to the VM entity vertex
- Inspect: VM has been in deleted state for months, but the libvirt, Cephimage and OVS states still remain.
- Possible bug in retrying / recovery mechanism

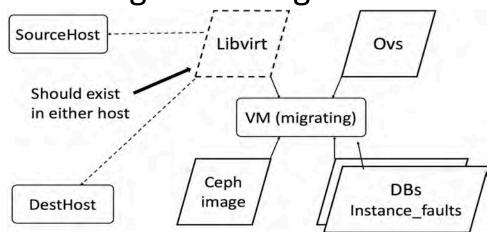




Database mismatch case

Anomaly case study: failed VM migration

- Subgraph: the migrating VM is missing libvirt state, both from the source host and the destination host
- Inspect:
 - Database: nova.instance_faults shows "cannot remove config /etc/libvirt/..." on the source, and "error removing image" on the destination
 - Log: "instance not resizing, skipping migration" repeated 653 times
- Possible resource management bug



Future work

- Analyzing event and state history
 - Indicate **why** the system end up in an inconsistent state
 - Might predict the failure before actually happens



Future work

- Including other data sources
 - i.e. source code, bug reports and documentations
 - Help further **confirm** the proposed anomalies/bugs
 - Provide insights in how to **fix** the bugs discovered

E C Canonic	all Grove Ltd (GB), https://taunchp	ad.net/reutron	c	docs.openstack.org	۴	C Grikub, Inc. (US) https://github.co	om/openstack/openstack		R 0
			Documentation for Mitaka (April 2016) This is the Made release. Use the menu to select a prior release if needed.		Personal Open source B	Rusiness Explore Pricing Bo	g Support This imposition	Sign in Server	
Overview Code Bugs	Overview Code Bugs Blueprints Translations Answers								
Registered 2511-05-10 to 🗸 Re	IX CAR					openstack / openstack		@ Watch 459	# Star 1.579 ¥ Fork 888
Neutron is a virtual n It was previously kno	network service for Openstack. own as 'Quantum'.					O Code Pull requests 8	- Pulsa Graphs		
Just Like Opertitisch heve provides in API to dynamically request and configure virtual servers. Nature and oxide in API to dynamically reposit and configure virtual reports. These returns content: "Warfunder" in a sub-operational service (a.g., with Chino Nature Vind). The Nature API supports terminism to provide advanced network topoloticities (a.g., QOS, ACLS, network monitoring, st.)			C Release Notes	Subser G	http://spenstack.org				
			New features, upgrade and deprecation notes, known issues, and bug		④ 91,189 commits)/ 3 branchés	D 1 volence	II. 1,295 contributors	
	For submitting blueprints/RFII and defect reports please follow instructions at: http://			fixes End User Gr	End User Guit	Branch: master + Ame city measured			Find file Ones or comment in
	docs.openstack.org/developer/newtron/policies/blueprints.html For release informiation and deliverables, please also look at: http://releases.openstack.org/				Create and may command-line	District station + American accession	Find file Commencement in		
Project information Series and miles			OpenStack Documentation Release Notes Command	Comment I in	Cooph committed with openatack-genit Uponted spinistical/openations			Latest convrid H12x34 & Hours ago	
				Comprehensive	Its activity-board @ 954ct59	Updated cowstack/operatack		4 months ago	
Part of:		R R ten		Getting stanted with the most commonly used OperStack services	Open source : Resources for a	Ta api-site @ 8526640	Updeleti operstack/operstack		8 hours ago
OpenStack	S fuestion Drivers					III api-wg @ 5d840ab	Updated operatack/operatack		27 diiyo iigo
Also known as: quantum	Orlean S Neutron Drivers					Ta askbot-theme @ 0cc9f5/	Updated openatack/upenatack		a year ago
Licence:	n Uberty-1 Ube	ten in the				Ita barbican @ ce63381	Updated operatedy/operatedy		3 cieve ago.
Apache Ucence		liberty-1 libert Incubator-two				Ta barbican-specs @ 6/58e84	Lipdeled operetack/operetack		2 months ago
♣ RDF metadata		incubator		Installation Guide for openSUSE Leap 42.1 and SUSE Linux Enterprise Server 12 SP1	Project	III bashate @ 8361128	Updated openstack/openstack		19 8695 890

Future work

- Supporting incremental graph construction
 - Capture the continuous evolving of state and event in a online way

- Applying SOSG to other systems
 - i.e. big data frameworks
 - General applicability

Conclusions

- System operation knowledge can be automatically discovered with simple rules
- Entities and links are important knowledge in systems, and can be captured with a state graph (SOSG)
- Many potential applications
 - System state query as graph traversal
 - Anomaly detection
 - ...

Thank You



We are hiring: faculty members, postdocs in any CS field contact: weixu@tsinghua.edu.cn