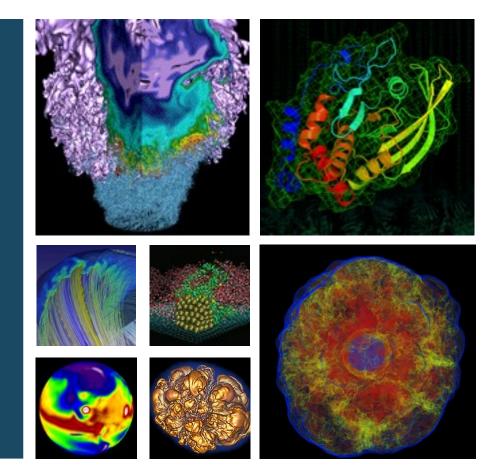
A slice of the NERSC data collect system





Thomas Davis & Cary Whitney ^{Title} ^{4/21/16}







- NERSC is moving between data centers.
 - Old data center is a mis-mash of systems.
 - 3 separate building control systems
 - One is Windows 98 based
 - Two are embedded Windows
 - This one is browser based, except..
 - You load an active-x object to check java version
 - And if you do not have the correct version, downloads it from a 3rd party website..
 - Then you load java to run a java-ws program..
 - Which then downloads and runs a Windows DLL to provide window decoration.
 - New data center has no chillers.







- Introduce several new ideas..
 - Clean slate no legacy to clean up or rebuild.
 - Security is paramount.
 - Most BAS/metering protocols are insecure.
 - Ability to Fail fast
 - Modularity
 - We demand the ability to replace anything, at any time, for any reason.
 - Use services that are designed for HA
 - We want failover, maintainability, and scalability
 - Zero downtime.
 - We are the first up, and the last down







- A new way of doing things.
 - System must be first one up, and the last one down.
 - System has many sensors, collectors and networking.
 - 1600+ temperature sensors
 - 800+ power meters
 - Substations master breakers, Substation breakers, panel breakers
 - 600+ power strips (Raritan, 50/60 amp is the preferred vendor/model)
 - With breaker level monitoring 6 breakers per strip.
 - 200+ bacnet devices
 - Many with logical sub-devices.
 - 2k+ ethernet ports
 - 24+ network switches
 - 4 Seismometers
 - The data roach hotel.
 - Data checks in, but never checks out..
 - HPSS is your friend.





Footer Information

- 4 -



- No plugins.
 - No java, no active-x, no flash
- HTML5 support.
- Desired to work on Cell phones, tablets, Linux, OSX, Windows.
- Stream processing
 - Ability to calculate a PUE in realtime
 - Fault detection
 - Security incidents
 - Building Control
 - Filtering
- Instrument the data collectors themselves.





Footer Information

- 5 -



Metrics 2.0 implementation

- What the heck is metrics 2.0?
- A different way of representing data
- Attaching meta information to the metric.
- See http://metrics20.org

Traditional systems	Metrics 2.0
<pre>collectd.dfs1.df.srv-node-dfs10.df-complex.used diskspacesrv_node_dfs10.byte_used { server: dfs1 }</pre>	<pre>{ server: dfs1 what: diskspace mountpoint: srv/node/dfs10 unit: B type: used metric_type: gauge } meta: { agent: diamond, processed_by: statsd2 }</pre>





- Ovirt (RHEL) VM management
 - VLAN tagging/trunks
 - PXEboot needs untagged packets..
 - Dense servers 8 nodes in 4U slot.
 - Supermicro Fat-twin based
- Rancher with Docker containers
- Split networks
 - Have to connect to 10 different networks.
 - 3 public networks
 - 7 private networks

Office of

Science

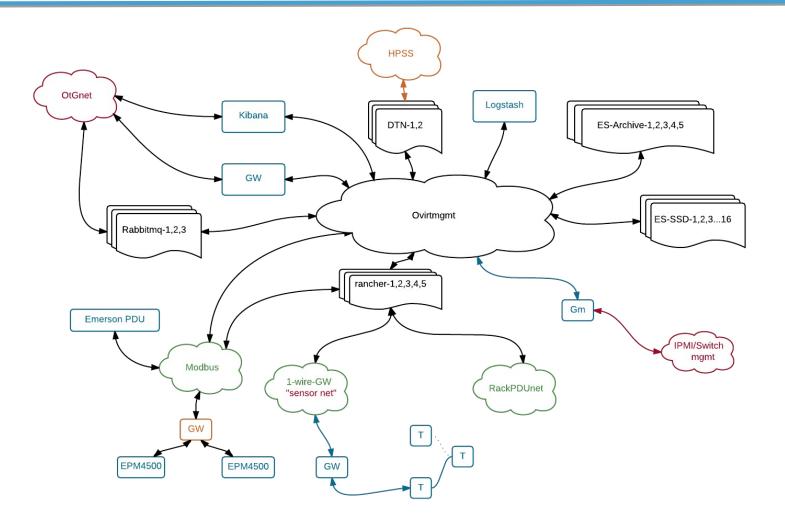






CRT Data Collect Network

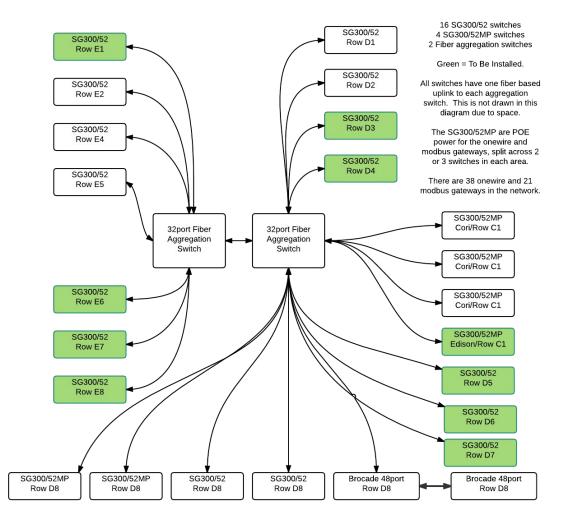










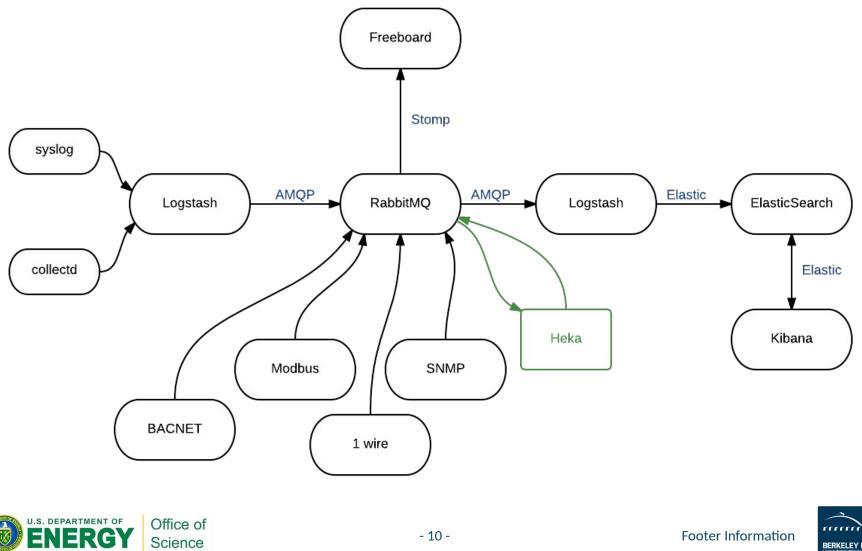






Science









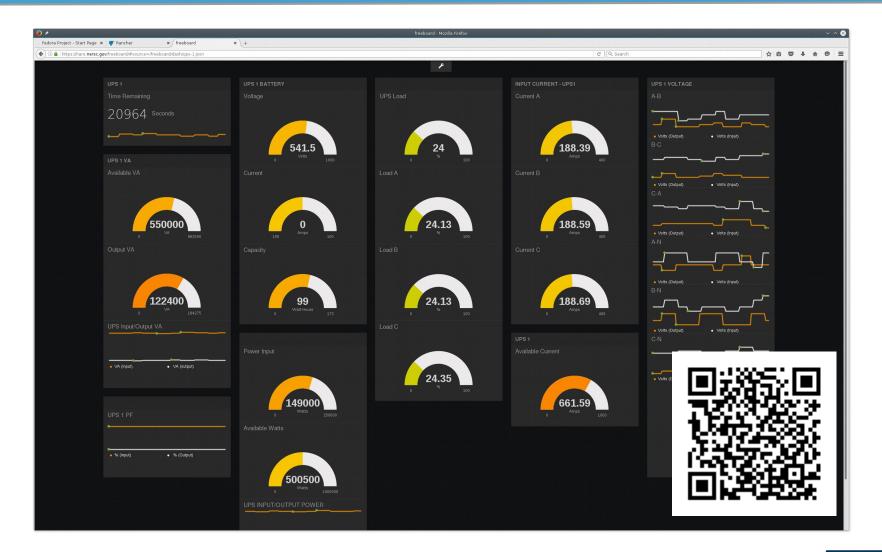


- 15 Elasticsearch (ingest) nodes w/SSD and 64GB ram
- 5 Elasticsearch (archive) nodes w/8TB disk and 64GB ram
- 3 dedicated Master nodes (both Consul & Elasticsearch) (vm)
- 2 internal logstash2amqp nodes (ha) (vm)
- 1 Kibana/Elasticsearch worker node (no data/not a master) (vm)
- 4 internal amqp2logstash nodes (vm)
- 3 rabbitmq-servers in ha/mirror cluster mode (vm)
- 6 node redis cluster (vm)
- 8 Rancher/docker nodes
 - 1 VM for Substation network
 - 2 VM for UPS network
 - 2 real nodes (dual socket, 32 cores, 64GB of ram)
 - 3 floating VM's











.....

BERKELEY LAB

oVirt



- http://ovirt.org
- Redhat's RHEV product, opensource version.
 - We are currently on the 3.6.x version.

				aVirt Engine Web Administrat					~ ~ @											
edora Project - Start Page 🛪 🕌 Rabi				👩 Grafana - env collectid 🗙 🔿 O oʻvi	IT Engrie Velo AL. X (+	@ Q. Search		****												
VIRT OPEN VIRTUALIZATIO			inta jino inti			i e j[e said		Igure Guide About F												
VITE OPEN VIRTUALIZATIO	N MANAGER								eesteloo											
ost:					◎ ×				irt Engine Web Administration - Mo		and the second second second second second	and the second		~ ~ 😒						
	Data Centers	Clusters Hosts	Networks Storage	Disks Virtual Machines	Fedora Project - Start Page 🛪 🏪 Rabi			nv] 🛪 🧑 Orafana -	erw collected */ O aviet Engine	/ w										
stem	New Edt R	errove Activate Maintenance	Select as SPM NUMA Support	Approve Reinstal Upgrade Configur	🔄 🔄 🗎 Https://hitter.nemsc.gov/ov/rt-er		ale=en_US#ims				C Q Search		2 E U	+ + ⊕ ≡						
and Al Collapse Al		Name C	emment Hostname/P	A Cluster	OVIT OPEN VIRTUALIZATION	N MANAGER					Fedora Project - Start Page 🗴 🕌 Rabi	Windowsenant y III and	memorial collect w keel	Retread T Co	ovit Engine Web Admin					
System	AL 17	d8-r13-c1-r8	d8-r13-c1-ri8	D8-R13-glusterfs	Vms:						(Carl Construction of the second seco				and encounter a) () ever trident mits to		e (9.		****
Data Centers		d8-r13-c2-r8	d8-r13-c2-n8	D8-R13-glusterfs	(Data Centers Clusters	Hosts Networks	Storage Disks	Virtual Machines Pos	ls Templa	0		internation of the second s						Narch	
T CRT	41 37	d9-r13-c3-r/8	d8-r13-c3-n8	D8-R13-glusterfs			VE Clane VM Run Once A			ico Make Terro	OVIRT OPEN VIRTUALIZATIO	MANAGER								
 Storage 	41 5	d8-r13-c4-r8	d8-r13-c4-n8	D8-R13-glusterfs	System	New VM Import Eat Herro	ve clane VM Hun Once .A.				Volumes:						×	* Q		
Networks	AI 🕅	d8-r13-c5-n8	d8-r13-c5-n8	D8-R13-glusterfs	Expand All Collapse All	Name	Comment Host	IP Address	FQDN	Cluster		Data Centers Cluste	rs Hosts Netwo	orks Storage Dis	ks Virtual Machine	es Pools	Templates Volum	Illeare		
C Templates		d8-r13-c2-n5	d8-r13-c2-r/5	D8-R13-LBL	(¥ 🕤 System	🔺 ! 📾 amqp2elastic-01	d8-r13-c3-n		amgp2elastic-01.crt.nersc							10018	Tempines Tempines	0.011		
Clusters		d8-r13-c3-r5	d8-r13-c3-n5	D8-R13-LBL	🗸 🔻 関 Data Centers	amqp2elastic-02	d8-r13-c1-n		amqp2elastic-02.crt.nersc			New Remove Start Str	p Rebalance Optimize for Vir	int Store Profiling + Snapshol	Geo-replcation +					51×
External Providers		d8-r13-c1-n1	d8-r13-c1-n1	D8-r13-net211	V CRT	A I amqp2elastic=03	d8-r13-c3-r0				Expand Al Collapse Al	Name	Cluster	Volume Type	Bricks	info	Space Used	Activities	No of snapshots	
Iocal-glance-image-reposition in the second seco	a 🔺 🗍 📅				 Storage 	amqp2elastic-04	d8-r13-c3-n	192.168.85.230	amgp2elastic-04.crt.nersc			🔺 ec0	D8-R13-glusterfs	Distributed Disperse	8 7 0		20%		0	
Iocal-glance-image-reposi		d8-r13-c1-n3	d8-r13-c1-n3	D8-r13-net211	Networks	▼ ! m bacnet				D8-R13-LE		▲ gi0	D8-R13-glusterfs	Replicate	▲3▼0		1196		0	
ovirt-image-repository		d8-r13-c1-n4	d8-r13-c1-n4	D8-r13-net211	Templates	🔺 ! 📾 build?	d8-r13-c1-ri	192.168.85.233	build7.crt.nersc.gov	D8-r13-net		🔺 ovirt	D8-R13-glusterfs	Replicate	📥 3 🕶 0		11%		0	
Errata Guest Information		d8+r13-c2+r1 d8+r13-c2+r2	d8-r13-c2-n1 d8-r13-c2-n2	D8-r13-net211 D8-r13-net211	Clusters	🔻 💼 CentOS6				D8-r13-net	Storage	 ovirt-ssd 	D8-R13-glusterfs	Disperse	▲3▼0		21%		0	
Guest information		d8-r13-c2-n2	d8-r13-c2-r0 d8-r13-c2-r0	D8-r13-ret211	V AEdernal Providers	CentOS7				D8-r13-net	Networks	 scratch 	D8-R13-glusterfs	Distribute	▲7 ▼ 0		29%		0	
		d8+r13-c2-rt3	d8-r13-c2-n3 d8-r13-c2-n4	D8-r13-ret211	local-glance-image-reposi		d8-r13-c1-n		curator.ort.nersc.gov	D8-r13-net	Templates	▲ ssd	D8-R13-glusterfs	Distribute	4 🕶 0		0%		0	
		d8-r13-c3-r1	d8-r13-c3-n1	D8-r13-net211	local-glance-image-reposi		d8-r13-c3-ri d8-r13-c1-ri		docker-dex.ort.nersc.gov											
		d8-r13-c3-r2	d8-r13-c3-r2	D8-r13-ret211	C ovirt-image-repository	 docker-registry env-ha-1 	d8-r13-c1-n d8-r13-c2-n		docker-registry.crt.nersc. env-ha-1.crt.nersc.gov	D8-r13-net										
					Guest Information		d8-r13-c2-ri d8-r13-c2-ri			D8-r13-net		10								
	General	Virtual Machines Networ	Interfaces Host Devices	Host Hooks Permissions	GOISt Intomation	env-ha-2	d8-r13-c2-rt		grafana.crt.nersc.gov	D6-R13-LE		×								
	3 4 11	Agrate Cancel Migration				a jan galara	d8-r13-c1-r0		hare.crt.nersc.gov	D8-r13-net										
	Name	Clust	er IP Add	fress FQDN		A le ipri	d8-r13-c2-n			D8-r13-net	Errata									
	a build7	D843	3-met211 192.10	68.85.233 build7		kibana	d8-r13-c3-ri		kibana.ort.nersc.gov	D8-r13-net	Guest Information									
	A Inagios	📾 08-0	3-net211 192.10	68.85.246 192.168.95.204 nagio:	6	logstash2amop-01			logstash2amgp-01.crt.ner		1									
	A (rabbit-2	💼 08-r	3-met211 192.16	68.85.219 128.55.211.68 rabbit	6	logitash2amop-02			logstash2amqp-02.crt.ner											
okmarks						A Im master-1	d8-r13-c2-n		master-1.crt.nersc.gov	D8-r13-net										
js						master-2	d8-r13-c1-n		master-2.crt.nersc.gov	D8-r13-net										
Message: /pr 21, 2016 2:4	7:53 AM Dete	cted conflict in hook start-POS	-31ganesha-start.sh of Cluster	r D8-R13-glusterts.		A I m master-3	d8-r13-c2-ri	192.168.85.247	master-3.crt.nersc.gov	D8-r13-net	2									
						🔺 💼 motd	d8-r13-c2-rú	192.168.85.150 12	motd.ort.nersc.gov	D8-r13-net	2									
						A I m nagios	d8-r13-c1-ri	192.168.85.246 19.	nagios	D8-r13-net	2									
					Bookmarks	A I @ netdisco	d8-r13-c3-n	192.168.85.221 19.	netdisco.crt.nersc.gov	D8-r13-net	4									
					Tags	A I an nms	d8-r13-c2-rc	192.168.85.183 19.	nms.crt.nersc.gov	D8-r13-net	2									
					al Last Message: Apr 21, 2016 2:4					_										
					Apr 21, 2016 2:47:53 AM (x)															
											Bookmarks									
					Apr 21, 2016 2:47:53 AM 🛞						Tags									
					✓ Apr 21, 2016 2:45:41 AM											•	_			Avens (72) St Events
					✓ Apr 21, 2016 2:45:35 AM 🛞															New York (12)
					✓ Apr 21, 2016 2:44:40 AM x					onas, Actual de										
					✓ Apr 21, 2016 2:44:35 AM															
					✓ Apr 21, 2016 2:43:51 AM 🛞	Mgration completed (VM: red	is-c-5, Source: d3-r13-c1-n1,	Destruitor: d8-r13-c3-n2, D	uration: 17 seconds, Total: 17 s	reconds, Actual	 Apr 21, 2016 2:45:41 AM (x) 									
											🗸 Apr 21, 2016 2:45:35 AM 🛞									
											🧭 Apr 21, 2016 2:44:40 AM 🛛 😠									
											🖌 Apr 21, 2016 2:44:35 AM 🛛 🛞									
											Apr 21, 2016 2:43:51 AM (x)									



BERKELEY

oVirt



- KVM/qemu based
- GlusterFS, nfs, iscsi volume management
- Web interface
- Understands VLANs, bridges
- Can talk to Openstack components
- Manages images, templates for you.



RancherOS



• http://rancher.com

- Recently released 1.0.1
 - Oh no.. a x.0 release!
 - Does have some irritating bugs.
 - But it does solve several major problems!

ra Project - Start Page: # 🖉 Aarcher # 👌 4	Ranh	er - Mazila Fredes	0 0 + + 0 =	1						
Applications catalog		API HELP	(A)~	1	er. Maalla Parlas	x A B				
HOSTS CONTAINERS	STORAGE POOLS CERTIFICATES	REGISTRIES	Federa Project - Start Page ★) ♥ Rancher ★ \+ ♦) © 192 198 45 232 FERENCE at consists/consists/consists/		C (6, Seech	1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1				
Sts Add Host			APPLICATIONS CATALOG	INFRASTRUCTURE ADMIN [®]	API HELP	(%) v	. Randow - Muslin Frahm	_	_	_
			ALL LIBRARY NERSC			Tedera Tropet - Start Pape x € ▼ Another x ↓ 4 (♠) ① 102.268.85.202.00000vr/s1st/sppcotacles			0.0	
	ACTIVE II :					APPLICATIONS CATALOG INFRASTRUCTUR				
d8-r13-c5-n1	d8-r13-c5-n2	rancher-01	Catalog: NERSC					-		
entOS Linux 7 (with KVM) (3.10.0-327.10	Å CentOS Linux 7 (with KVM) [3.10.0-327.10	∆ CentOS Linux 7 (3.10.0-327.13.1.el7.x86_64)				STACKS				Enviro
32x24 GHz = 62.7 GB = 415 GB	B 32x24 GHz B 62.7 GB B 415 GB	4x2.4 GHz GHz 7.64 GIB 30.7 GIB net.modbus-true net.onewire-true	2	0,0	0,0	Stacks Add Stack			Sort By:	State Na
ack: logspout	Stack: logspout	Stack: logspout							-	
.logspout_1 10.42.222.59 [O.logspout_2 10.42.101.143 1	O.logspout_3 10.42.236.101	modbus	MotionEye	Onewire Collectors	+ logspout	Up to date Add Service 💙	1 Service C	8 Containers	0 :
cic modbus	Stack: modbus	Stack: onewire	Modbus sensor poller	MotionEye using motion software.	(Experimental) Onewire server, workers, and jobber				10	
r-modbus_pdu_5 172.17.0.7	O r-modbus_pdu_8 172.17.08	O .owserver_1 None 1			workers, and jobber	modbus	Up to date 🛛 Add Service 💙	Services (Containers	0 :
D.master_1 10.42.118.240	O r-modbus_master_1 172.17.0.6	O .worker_5 10.42.67.102	View Details	View Details	View Details	@Active ion ①	Image: docker-	Service	2 Containers	0 :
.pdu_1 10.42.33.217 1	Stack: onewire	O .worker_10 10.42.181.130					registry.crt.nersc.gov:5000/pulsar_worker_1			
.pdu_11 10.42.206.60	O worker.4 10.42.41.149	O .worker_15 10.42.77.17				⊕ Active master ①	Image: docker- registry.crt.nersc.gov:5000/pulsar_worker_1	Service	1 Container	0 :
.pdu_12 10.42.203.49	O.worker_9 10.42.112.160	Standalone Containers	0,0			⊕Active pdu ①	Image: docker-	Service :	5 Containers	0 :
D.pdu_13 10.42.224.206	O.worker_14 10.42.248.99	O Network Agent 10.42.79.89					registry.crt.nersc.gov:5000/pulsar_worker_1			
D.pdu_14 10.42.185.146	Standalone Containers	+ Add Container	-			Active ups ①	Image: docker- registry.crt.nersc.gov:5000/pulsar_worker_1	Service 2	2 Containers	0 1
indalone Containers	O Network Agent 10.42.128.74		RQ Web Dashboard							
Network Agent 10.42.31.106	+ Add Container	ACTIVE II : ups-gw-1	Web interface to peek into the RQ queues.			onewire	Up to date Add Service 💙	3 Services (21 Containers	0 :
+ Add Container	ACTIVE	 ⁰ 192.168.85.151 ⁴ 1.10.3 [∆] CentOS Linux 7 (3.10.0-327.13.1.el7.x86_64) [−] 	View Details			@Active jobber ①	Image: docker-registry.crt.nersc.gov:5000/mini/owfs- jobber:0.9	Service	1 Container	0 :
						di Active owserver ()	Image: docker-registry.crt.nersc.gov:5000/mini/owfs- server:0.18	Service	1 Container	0 :
						Active worker ()	Image: docker-registry.crt.nersc.gov:5000/mini/owts- worker:0.7	Service 1	19 Containers	0 1
						+ rq-dashboard	Up to date Add Service 🗸	1 Service	1	0 :





Rancher



- Supports docker, kubernetes, and Docker Swarm.
- Internal, encrypted network between containers.
- HTML and CLI interfaces.
 - Docker compose file support
- Catalog for applications.
- Running on both VM's and bare metal systems.
- They even have a Docker based OS RancherOS
 - Like CoreOS/AtomicOS, but geared towards virtual machines.
- Convoy for docker volume management using GlusterFS/NFS.





Rancher



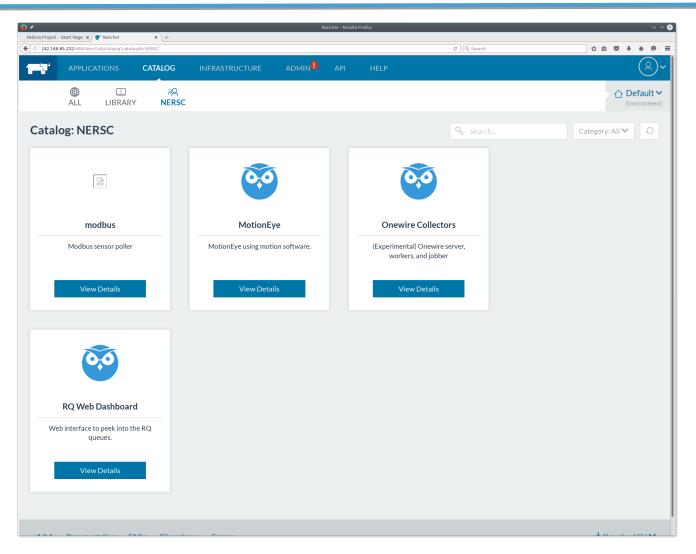
- One click deployment of application stack.
 - Scales containers, creates internal network, manages volumes
 - Rollback, upgrades are simplified.
 - Schedules containers on the proper nodes.
 - Based on tags.
 - We use this for network routing for example.





Catalog..









Setup.



) *		_	Rancher - Mozi			~ ^ 😣
Fedora Project - Start Page 🗴 💾 RabbitMQ Managen		collect x kopf[crt-env]	🛪 🧑 Grafana -	env collectd 🗙 🛛 o Virt Engine Web Ad 🧿		
(i) 192.168.85.232:8080/env/1a5/catalog/NERSC:on				e	l, Search	
APPLICATIONS CATALOG	INFRASTRUCTURE	ADMIN ^{ID} API HE	_P			([®])~
ALL LIBRARY NERSC						Crivironment
Catalog: Onewire Collectors *						
Catalog: NERSC Category: Onewire Maintainer: Thomas Davis <tadavis@bl.gov></tadavis@bl.gov>		Onewire Collec				
Template Version						
0.1			-			
Select a version of the template to deploy						
New Stack						
Name			C	Description		
onewire				Description		
Configuration Options						
Worker Nodes:*						
1						(N)
Number of workers for collection						
Start services after creating						
			PREVIEW			
			Launch	Cancel		
v101 Documentation EAOs Eilean Iss	io Forume					Download CLLM





Production..



	e 🗙 📙 RabbitMQ Management : 80/env/1a5/apps/stacks	III environmental collect	× kopf[crt-en	10]	x 😤 Grafana - env collectd x O oVirt Engine Web Ad x freeboard	× 🐺 Rancher	× 公自 💟	
APPLICA		INFRASTRUCTURE		API	HELP			• "
STACKS								De Env
acks Add St	ack						Sort By:	State
	spout				Up to date Add Service 💙	1 Service	8 Containers	0
dl Active	logspout ()				Image: rancher/logspout-logstash:v0.2.0	Service	8 Containers	0:
⊗ – mo	odbus				Up to date Add Service 💙	4 Services	10 Containers	0
⊕ Active	ion 🛈				Image: docker-registry.crt.nersc.gov:5000/pulsar_worker_1	Service	2 Containers	0:
dł Active	master (j)				Image: docker-registry.crt.nersc.gov:5000/pulsar_worker_1	Service	1 Container	0
- Active	pdu ①				Image: docker-registry.crt.nersc.gov:5000/pulsar_worker_1	Service	5 Containers	0
⊕ Active	ups 🕕				Image: docker-registry.crt.nersc.gov:5000/pulsar_worker_1	Service	2 Containers	0:
	ewire				Up to date Add Service 💙	3 Services	21 Containers	0
- Active	jobber (j)				Image: docker-registry.crt.nersc.gov:5000/mini/owfs-jobber:0.9	Service	1 Container	0:
& Active	owserver (j)				Image: docker-registry.crt.nersc.gov:5000/mini/owfs-server:0.18	Service	1 Container	0 :
dłł Active	worker ①				Image: docker-registry.crt.nersc.gov:5000/mini/owfs-worker:0.7	Service	19 Containers	0 :
⊗ – rq-	dashboard				Up to date Add Service 💙	1 Service	1 Container	0
₼ Active	rq-dashboard 🛈				Image: docker-registry.crt.nersc.gov:5000/mini/rq-dashboard:0.1 Ports: 9181	Service	1 Container	0 :







- Service discovery/registration with Consul
 - Uses Nagios checks for health and integration in nagios.
 - DNS routing and load balancing
- Used by onewire stack to route onewire api requests to the proper gateway.
- Nagios also checks for service health.
- Also used to provide round-robin DNS name service for internal services.
- Future uses include configuration templating and key value pairs.



