Easily Extending a Usable z/VM System!

Vic Cross
Senior Technical Specialist
IBM zAcceleration Team

viccross@au1.ibm.com

Contents

Introductions: Vic Cross

Long time Linux user
- Multiple platforms
- Multiple distributions

Long time system administrator
- Linux, OS/2, Mac OS (macOS), Windows
- MVS, OS/390, z/OS
- VM/ESA, z/VM
- AIX

User of exotic Linux distributions
- Debian
- Fedora
- Gentoo

Author and co-author of IBM Redbooks and other docs

Introductions: z/VM ESI

- Integrated system based on z/VM

- Includes pre-activated and -configured:
  - DirMaint
  - Performance Toolkit
  - RACF
  - Operations Manager
  - z/VM LDAP

- Linux guest with add-ons
  - The subject of this session!

Express *System* Install

A truly integrated **system**

Setup and operation of the ELAN relies on the known setup of the underlying z/VM

Examples:

- z/VM LDAP with SDBM
  - For RACF updates
- Operations Manager
  - For Linux console interaction
- Prepared user IDs, defined to LDAP LDBM
  - For LDAP authentication integration
- Tools/utilities installed in z/VM
  - Used as part of the Linux-based automation

What is an ELAN?

By Grenadille own work, file link, CC BY-SA 4.0

By Brian Snelson from Hockney, Essex, England file link, CC BY-SA 4.0

By nakhon100, file link, CC BY 2.0

By SG2012, file link, CC BY 2.0

What is an ELAN?

Express
Linux
Automation and
Networking

“Run your z/VM with ELAN*!”* ☺

Pre-installed virtual machine running Alma Linux

- Web interface
- Support infrastructure for RHOCP:
  - DNS server (BIND)
  - Load balancer (HA-Proxy)
  - HTTP(S) proxy (Squid)
  - IP Masquerade
  - NTP server (chronyd)

Platform for our Ansible automation

ELAN runs Alma Linux

Initially we used RHEL

Potential issues here:

- Entitlement
- Maintainability (updates)

However, it’s “blessed” in most corporates

Alma is a community re-spin of RHEL

Binary compatible with RHEL

- Familiar operating environment

It is supportable, no $$$ required

No “blessing”

Cards and letters, please!

What does the ELAN provide?

Web interface

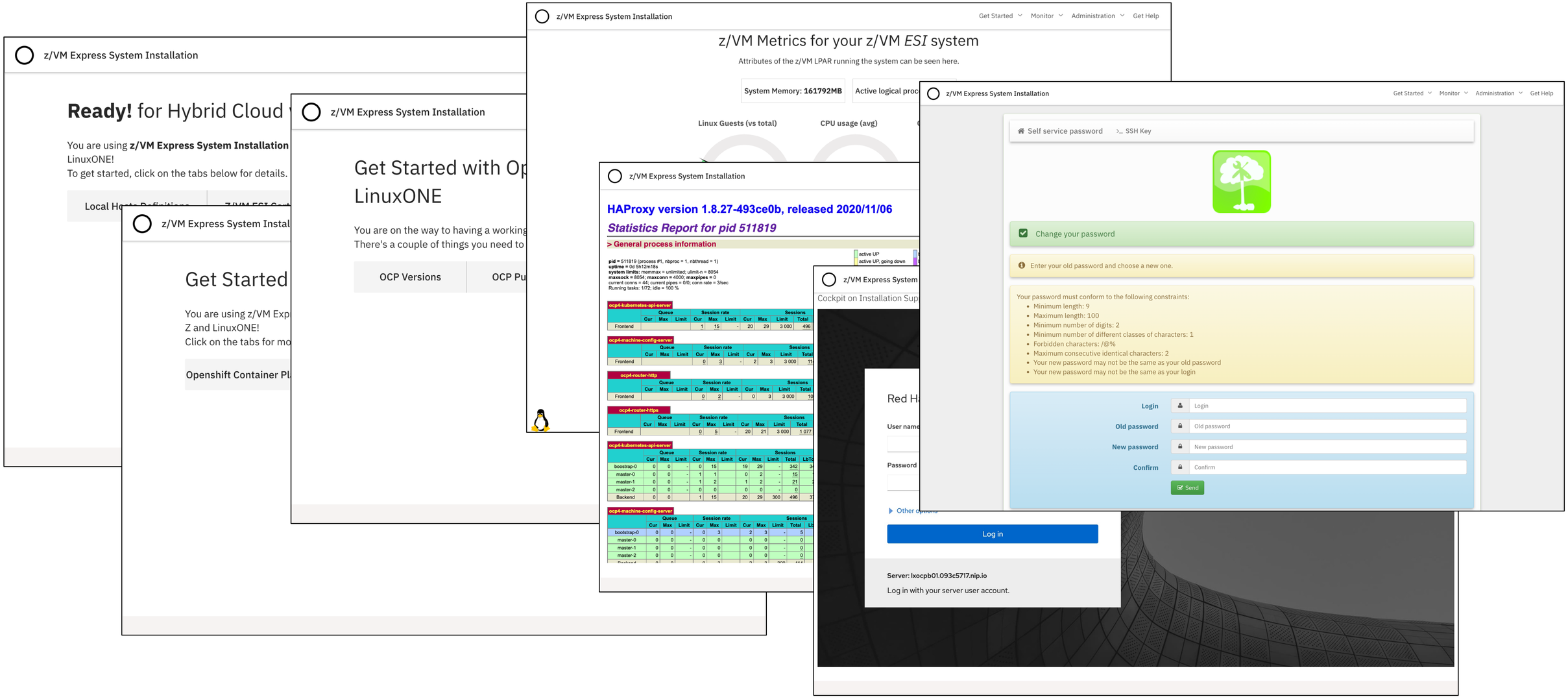
- Front-end for automation
- Simple z/VM metrics display (dials)
- LDAP/RACF utilities
- Access to other web views:
  - HA-Proxy statistics
  - Performance Toolkit
  - Cockpit (Linux web management)
- Launcher for RHOCP console

Extras

Automation to deploy
- RHOCP cluster
- IBM Cloud Infrastructure Center

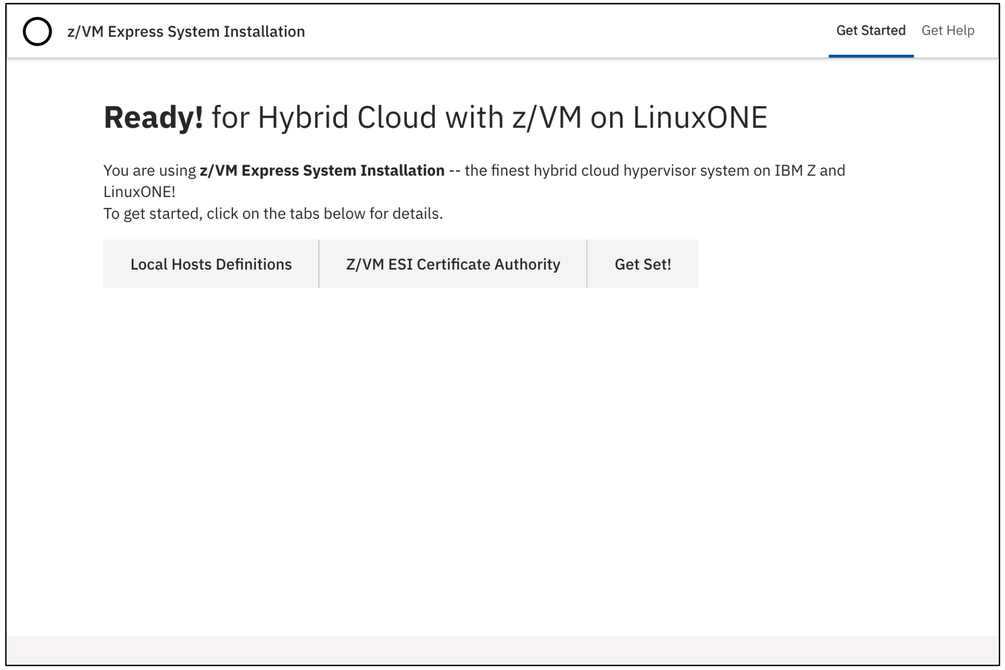
Plans
- Multi-system RHOCP cluster
- Multiple RHOCP clusters
- “Modularised” content
- Disconnected RHOCP installation
- Reduction of tasks requiring 3270 access

ELAN Web interface



Initial page is non-secure

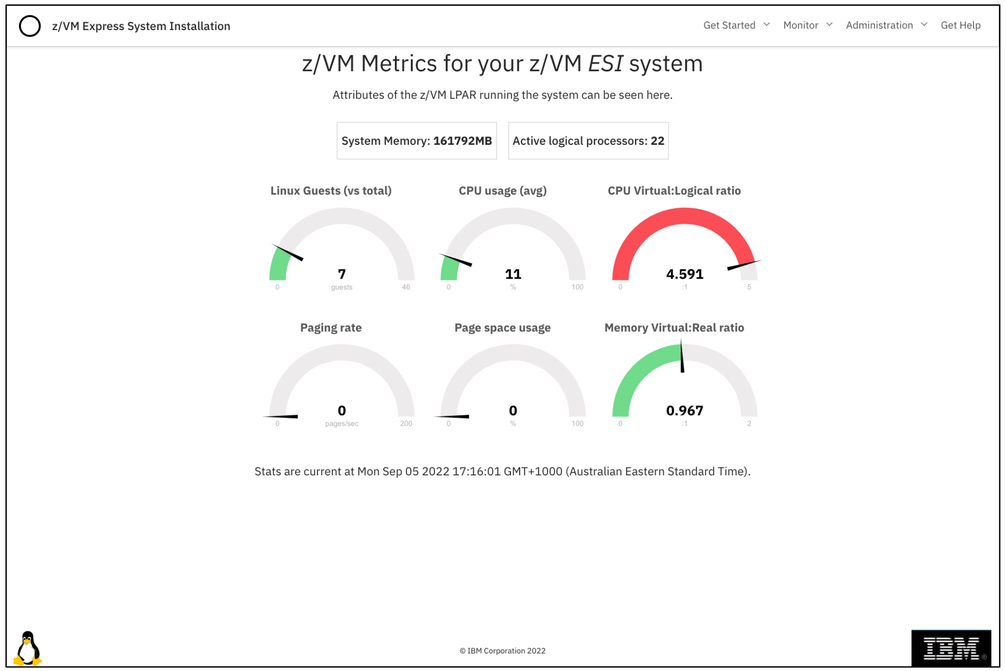
Design point: avoid browser security exceptions!



- Certificates for all secure interfaces signed by our own CA
  - ELAN web interface
  - z/VM: TN3270, Performance Toolkit, LDAP
  - Linux displays: HA-Proxy, Cockpit
  - RHOCP and ICIC consoles (if used)
- Non-secure initial page gives instructions for how to add the CA certificates to your trust store
- Things still work (mostly) if you can’t add the CA

Metrics display

Sometimes we just want to see simple dials!

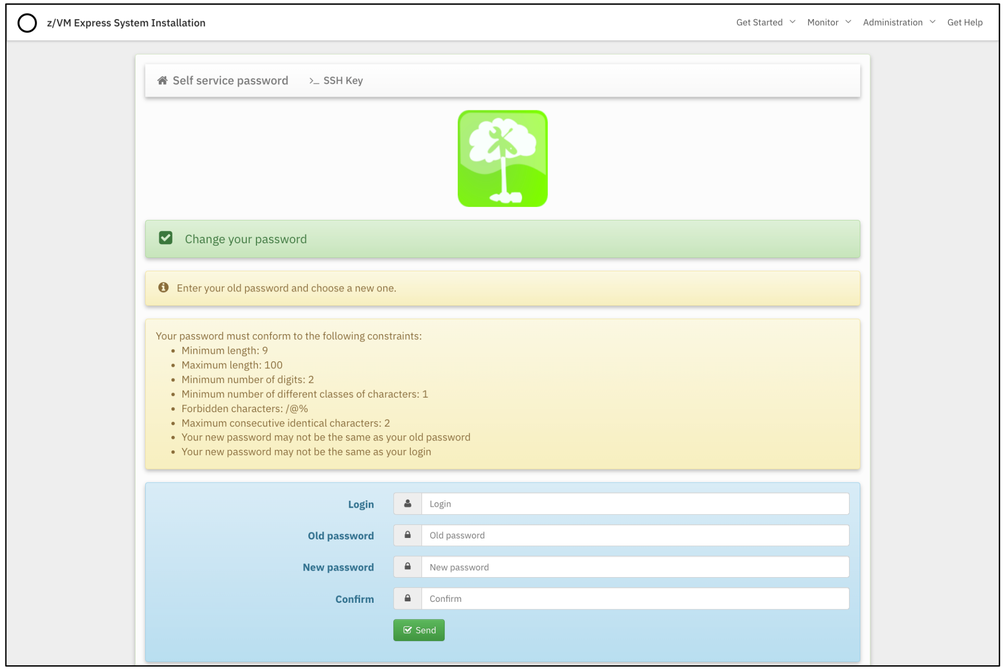


- Generated by Perl script parsing output from some CP commands
  - `INDICATE` (USER, NSS, LOAD)
  - `QUERY` (NAMES, USER, VDISK, PROC, STOR)
  - `QUERY ALLOC` (SPOOL and PAGE)
- Writes data to a JSON file, which the web page fetches using XHR to feed to the dial-draw code

- Working on some simple recommendations based on interpretation
  - Warn if “too much” CPU consumption by non-Vh processors
  - Warn if total guest defined memory > central + page space
    - If there is standby memory, adds a reminder that a memory re-config could address

z/VM LDAP

(Red Hat dropped OpenLDAP from RHEL 8… but we haven't missed it! ☺)



- LDBM and SDBM backends
- LDBM is configured as an authentication source for RHOCP and ICIC
  - LDBM uses Native Authentication → passwords in RACF
    - A web-based password change utility is provided
  - Demonstrates central ID management, without relying on that being present

- SDBM is used for a variety of purposes
  - "RACF Manager"
  - Within the automation for updating RACF profiles

ELAN Automation

Ansible augmented

- Most big work done with Ansible
  - RHOCP: "multiarch-ci-playbooks” was the foundation
- A **lot** of shell scripting
- Appearances also by **`expect`**
- No Ansible module(s) for z/VM
  - Used a Linux system and SMAPI
  - Idempotence…?
- LDAP SDBM was inspired! ☺
  - Avoided all the “how can I make this RACF change?” issues
  - Can’t use the Ansible LDAP modules though, due to object and schema expectations

RHOCP automation

Ansible playbook actions:

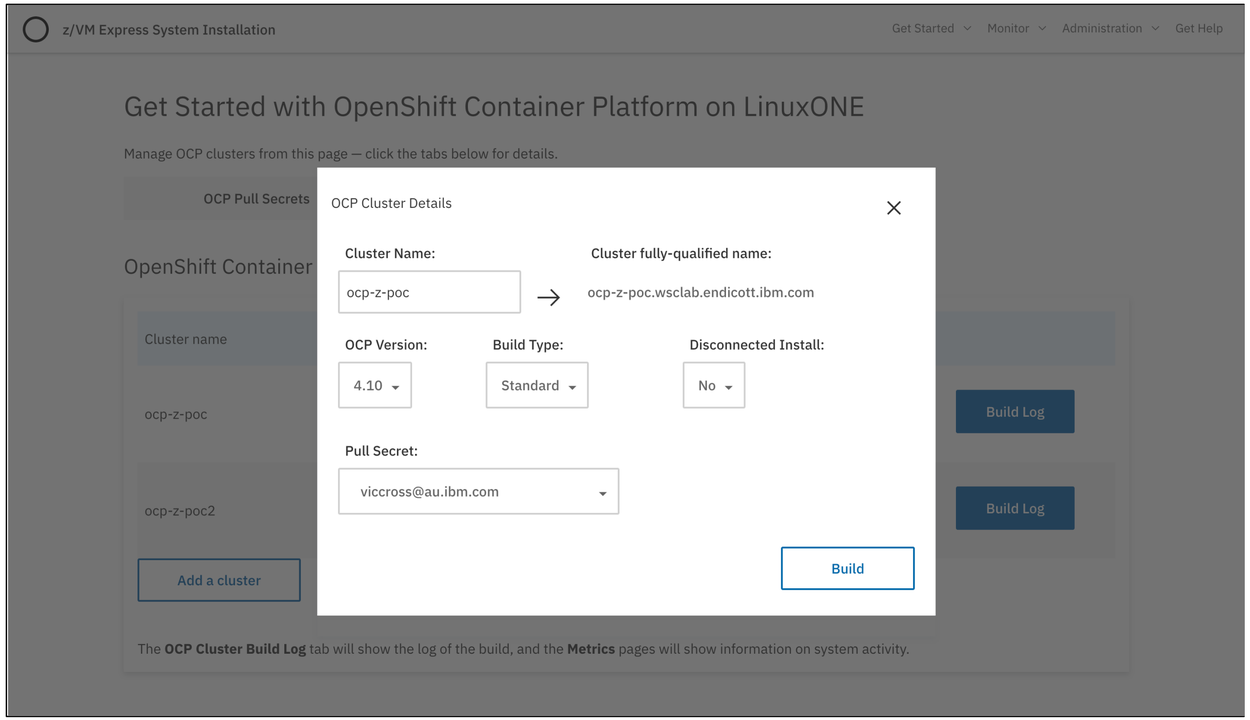
- Creates config files for boot loader
- Creates new ingress certificate
- Creates `install-config.yaml` from template
- Runs `openshift-install` to create Ignition files
- Injects NTP config into Ignition files
- Boots guests (via SMAPI)
- Waits for “bootstrap-complete”
- Shuts down bootstrap guest
- Renames it to third worker
- Boots third worker
- Waits for “install-complete”
- Changes cluster ingress certificate to the new one created earlier
- Configures LDAP authentication using z/VM LDAP

RHOCP automation – install-config.yaml template

```
apiVersion: v1
baseDomain: "{% raw %}{{ cluster_base_domain }}{% endraw %}"
proxy:
  httpProxy: http://{{ bastion_private_ip_address }}:3128
  httpsProxy: http://{{ bastion_private_ip_address }}:3128
  noProxy: .{% raw %}{{ cluster_name }}.{{ cluster_base_domain }}{% endraw %},169.254.169.254,{{ subnet_cidr }}
compute:
- hyperthreading: Enabled
  name: worker
  replicas: {{ cluster_nodes['workers'].keys() | length }}
controlPlane:
  hyperthreading: Enabled
  name: master
  replicas: {{ cluster_nodes['masters'].keys() | length }}
metadata:
  name: "{% raw %}{{ cluster_name }}{% endraw %}"
networking:
  clusterNetworks:
  - cidr: 10.128.0.0/14
    hostPrefix: 23
  networkType: OpenShiftSDN
  serviceNetwork:
  - 172.30.0.0/16
platform:
  none: {}
pullSecret: '{% raw %}{{ ocp4_pull_secret | to_json }}{% endraw %}'
sshKey: '{{ bastion_pubkey.content | b64decode | trim }}'
```

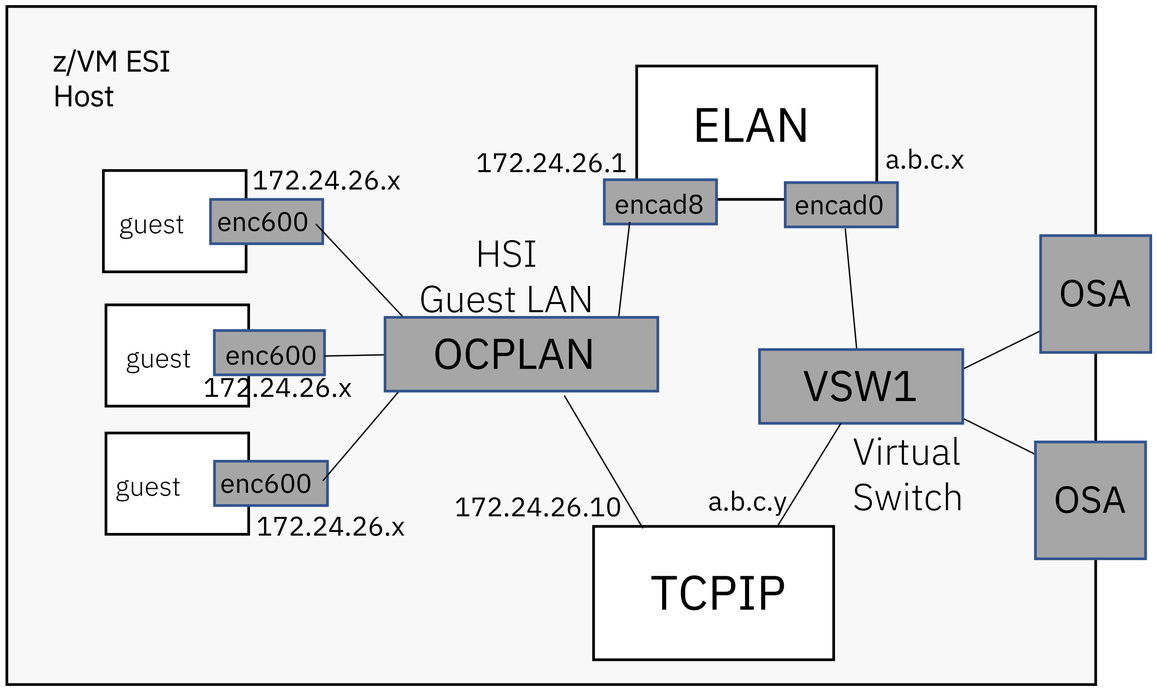
RHOCP automation

Futures…



- More customisability
  - Change the cluster name
  - Different cluster types
    - Three-node
    - Multi-system
    - Infrastructure nodes
  - Multiple clusters
    - DNS and HA-Proxy changes
- Disconnected installs
  - Aka “air-gap”, “mirror install”
- Networking
  - Multiple LPARs
  - RoCE

RHOCP automation – networking



- Cluster is on isolated network
  - Currently HiperSockets Guest LAN
- Great for PoCs
  - Fewer IP addresses needed
  - Everything works through the ELAN
- Support functions on the ELAN:
  - Reverse HTTP proxy (Squid)
  - NTP Server
  - IP Masquerade enabled
- Challenges for production use
  - Multi-LPAR clusters
- Automation to decide when the networking needs to change
  - IP address configuration as it happens now won’t be sufficient

ICIC automation

Simplifying deployment of your IaaS layer

- Different approach from RHOCP
  - RHOCP brings its own OS (CoreOS)
- Deliver ICIC without ballooning the static download size of ESI
- Automation comes in parts
  - Small RHEL image
  - ICIC prerequisite RHEL package repo
  - Small support files
- ICIC code obtained separately

ICIC automation – stages

1

Specify location of ESI files for ICIC deployment. May be the same location as ESI was installed from. If location is valid, deployment can be started.

2

Automation script starts:

- Creates guests
- Copies RHEL image and customizes (IP+hostname)
- Adds additional disk space to guests
- Boots management node guest

3

Inside management guest, script resizes root filesystem into extra disk space

4

Main script starts Ansible play:

- Waits for SSH
- Creates TLS certificate
- Creates ICIC properties file
- Runs the ICIC silent installation
- Replaces certificate

ICIC automation – stages

5

Main script sets OPERATIONS on ICIC compute node, then boots it

6

Inside compute guest, script resizes root filesystem into extra disk space

7

Main script starts Ansible play:

- Waits for SSH
- Runs ICIC script to add z/VM host to ICIC

8

Main script starts final play:

- Runs expect script to configure LDAP auth

ICIC automation

Improvements

- Configure network in ICIC
  - We have *most* of the required data
- Provide an ICIC-deployable Linux image
  - (or point folks to a source)
- Tie-in to “system discovery”
  - One ICIC management node per ESI group
  - Streamline multi-system ICIC deployment
    - If adding ICIC to an existing ESI group, deploy across all group members
    - If adding a new ESI member to a group with ICIC deployed, automatically deploy ICIC compute to the member

ELAN System Discovery

System discovery

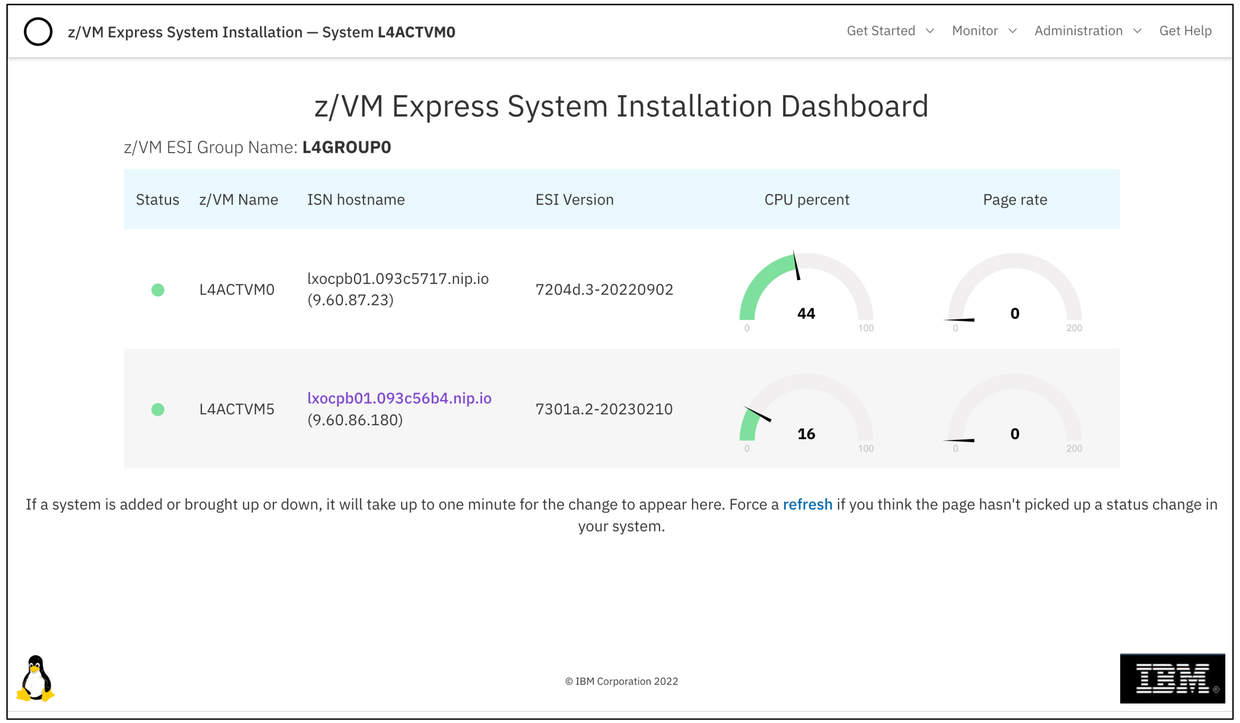
Growing beyond proofs-of-concept...

What does ESI need?

- ESI provides a best-practices z/VM
  - Decades of practical experience
  - z/VM management products
- Useful beyond just PoCs!
- Production RHOCP clusters...
  - Deployed across multiple z/VMs
  - Multiple nodes, infra nodes...
- How to make automation aware?
  - SSI is ECKD-only, which would preclude a lot of LinuxONE sites

Finna

Icelandic for
“to find”

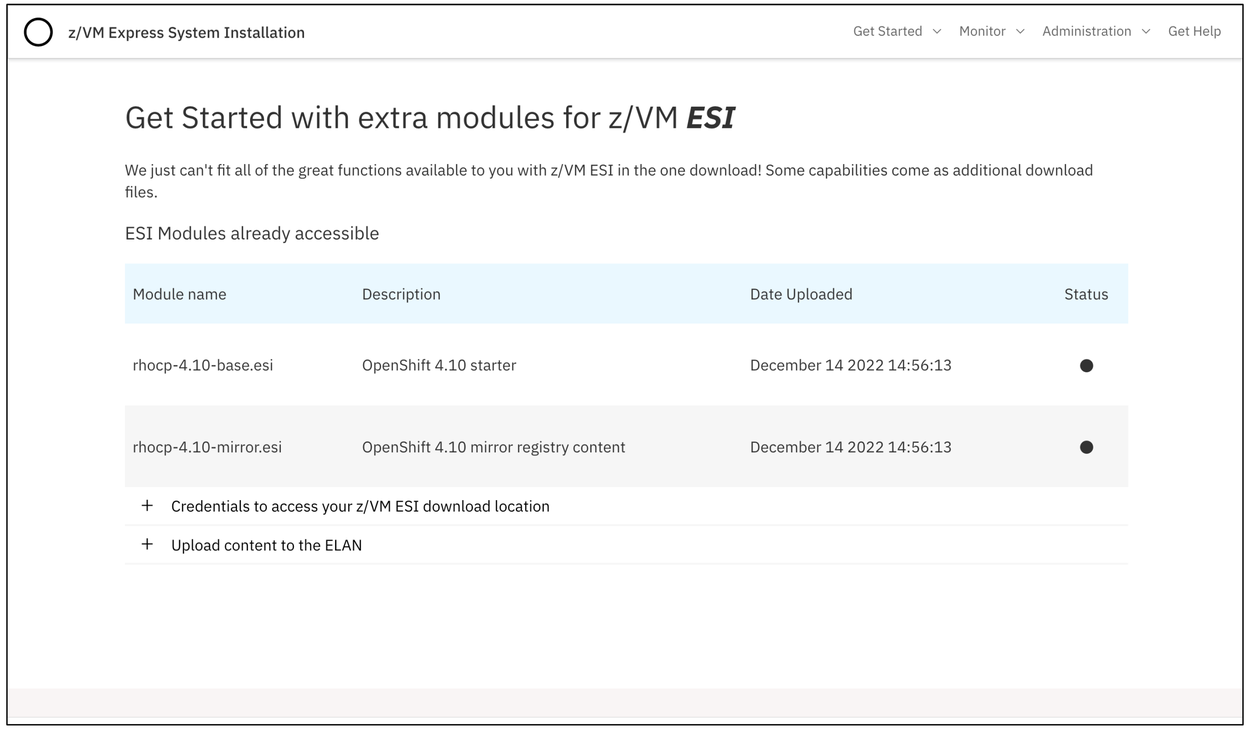


- Uses UDP multicast to find ESI systems on a subnet
  - Very simple messages
  - Based in principle on OSPF “hello” but much less sophisticated
- Command-line client to query
  - State of systems in a group
  - Details of systems (hostname, etc)
- Creates a dashboard of systems
  - Get a simple overview of ESI systems in the group

Extensibility

ESI Modules

Growing your z/VM system easily



- Components initially hard-built into ELAN
  - e.g. RHOCP installation support
- Inflexible! Adding functions would cause:
  - Bloat of the ELAN
  - Entirely new build of ESI to be made
- Add additional functions with “modules”
  - Download just what you need
    - RHOCP versions
    - RHOCP mirror registry content
    - …
- Over time, become a model that vendors could contribute to

Field updateability

Growing beyond proofs-of-concept...

What does ESI need?

- Using a supportable Linux is only the start
  - z/VM updates
  - ESI itself (ELAN code)
  - RHOCP
    - Especially if air-gapped

- Migrating the ELAN build from Ansible-managed file deposit to RPM-based *or* containerised
  - Update ELAN components separately from entire ESI system

- Modularising RHOCP content *should* help:
  - Download and enable a new RHOCP mirror blob
  - Inject some YAML, cluster sees available update

Thoughts…
Suggestions…
Ideas…

Thanks for Listening!

Our other sessions:

- ***Installing a Usable z/VM System is Easy!***
  - Friday 8:30 AM Traditions room
- ***Documenting, testing, and packaging of an automated bundle like z/VM ESI***
  - Saturday 9:45 AM Cartoon room

- Thanks to our Skunkworks team:
  - Ernie Horn
  - Bruce Hayden
  - Paul Novák
  - Vic Cross

- With help from:
  - Fred Bader
  - Jacob Emery
  - Jay Brenneman
  - Justice Heughan
  - Matt Mondics
  - Stephanie Rivero