# Distributing and Trusting Images between Cloud Providers

Loren Klingman

School of Computing

Clemson University

CLEMSON School of COMPUTING

# Abstract

- Computing is shifting toward cloud computing
- Clouds serve websites, offer scalable services, and power research at great speeds
- Users want customized environments
- Cloud providers must ensure security and policy compliance
- We build a chain of trust between image producers and cloud providers
- Endorsers verify images
- We extend the VMIC from CERN to export endorsed image catalogs and allow others to import the catalog

# Outline

- Background
- Related Work
- Technology Used
- Virtual Machine Image Catalog
- Conclusion

# Background

- Cloud Computing became a Web 2.0 Era Buzzword
- Customers want a Customized Environment
- Project built off of VMIC Developed at CERN
- We Implement External Trusted Image Providers

# Background - Grid vs Cloud

Grid Computing

- Project Oriented
- Interoperable
- Batch Scheduling
- Fast Network Connection
- Built to have various operating systems in one cluster
- Regulated by usage time

Cloud Computing

- All consumption monitored for payments
- Architecture virtualized
- Each system booted as available
- Usually boot virtual images so the user gets a custom environment on every system
- Not as interoperable

# Background - Why VMIC?

We want to verify that images are:

- Well Secured
- Not Malicious
  - Don't attack other sites
  - Don't attack our site using special permissions granted due to being behind the firewall
- Meet Our Standards

We also want to easily distribute house images and trust all images approved by certain other entities.

One example of a virtual machine policy:
http://www.jspg.org/wiki/Policy_Trusted_Virtual_Machines

# Background - VMIC Principles

- Basic Policy for Trusting Images
- Provide a distribution framework
- Flexible Machine Provisioning (within the pre-existing catalog of operating system and software configurations)
- Self managing in terms of distributing updated images
- Allow other catalogs to be trusted and imported
- Release metadata for image retrieval, verification,and identification (version, OS, software, etc)

# Background - VMIC Sharing Motivation

- Applies to both cloud and grid computing
- Benefits
    - Same image set across a variety of sites
    - Prevents data lock-in
- We add this to the VMIC from CERN
- Risks
    - Incorrectly shared images
    - Sharing insecure images
    - Having images changed in transit

# Background - Contributions

- Investigate trust and identity verification issues
- Propose a solution for endorsing and exporting an existing image catalog
- Prototype of the system

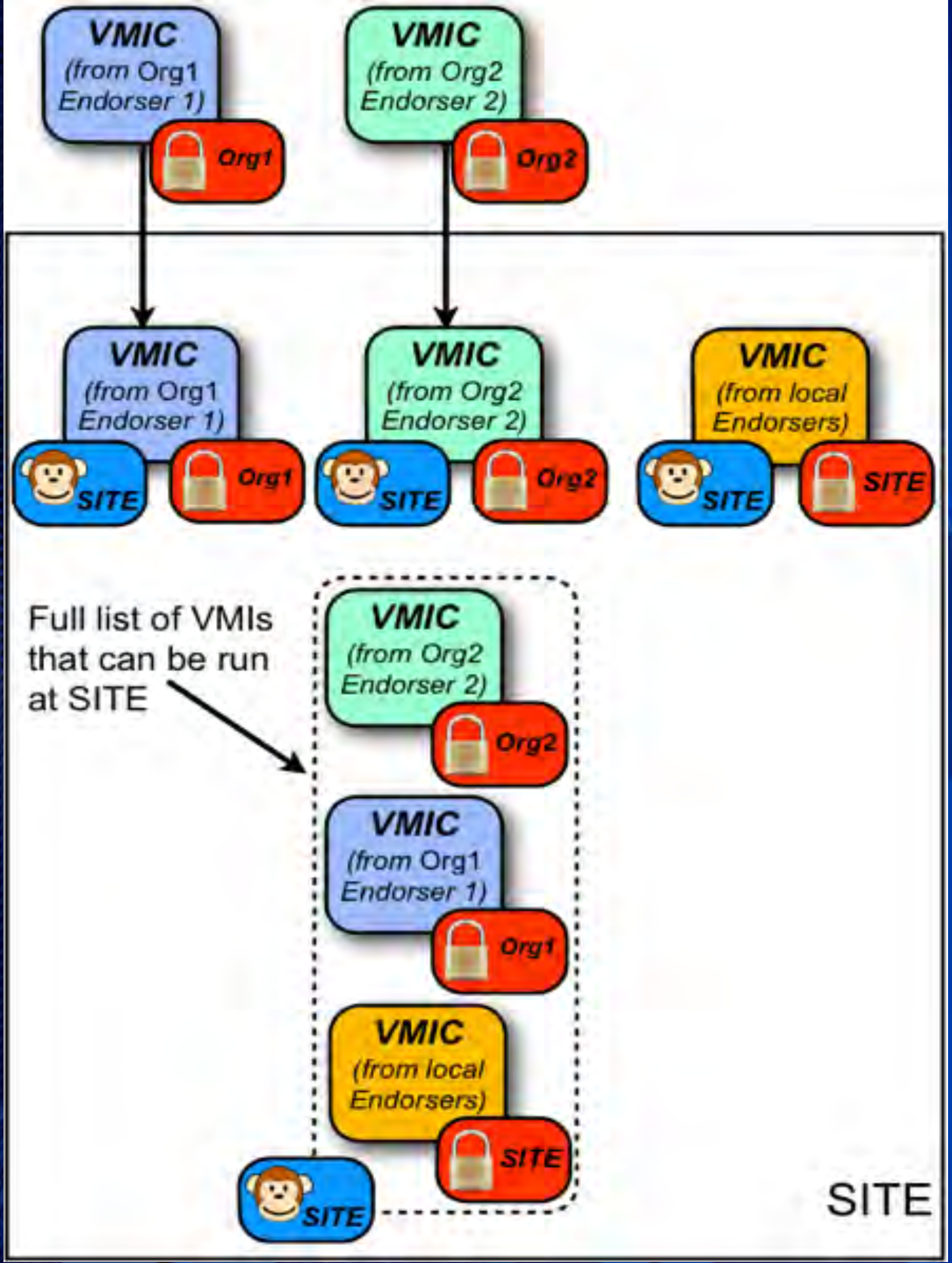
# Related Work

- Grid Computing
  - Sophisticated Data Processing
  - Research Computing
- Cloud Computing
  - Fifth Utility (water, electricity, gas, and phone)
  - Virtual Machines
  - Google AppEngine, Microsoft Azure, and Amazon EC2
- Virtual Machines
  - Developed by IBM in the late 1960's
  - Replicate operating system
  - Isolate programs, run new versions for testing
  - Work by Intel and AMD to optimize the hardware level

# Related Work

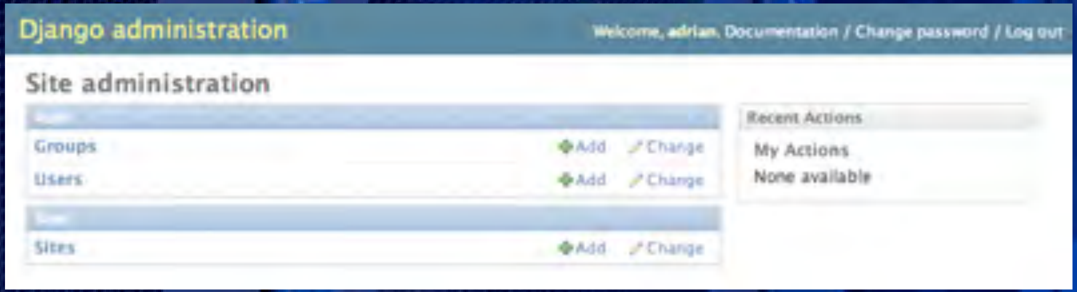
- Virtual Machines in the Grid
  - Support Legacy Applications
  - Layer of Security Between User Code and the System
  - Complete Environment Customization
  - Administrator Rights
- Virtual Machines in the Cloud
  - Similar Benefits as in the Grid
  - Research in provisioning, management/trust, weaknesses of Virtual Machine Hypervisors
- Virtual Machine Security Concerns
  - Virtual Machine Isolation
  - Securing Virtual Machines without knowing its state
  - See ACM Workshop on Cloud Computing Security

# Technology - VMIC



# Technology - Django

- High Level Python Framework
- DRY Principle (Don't Repeat Yourself)
- Easy to create a script with an administrator interface



Django administration

Welcome, adrian. Documentation / Change password / Log out

Site administration

| | | |
|---|---|---|
| Groups | Add | Change |
| Users | Add | Change |
| Sites | Add | Change |

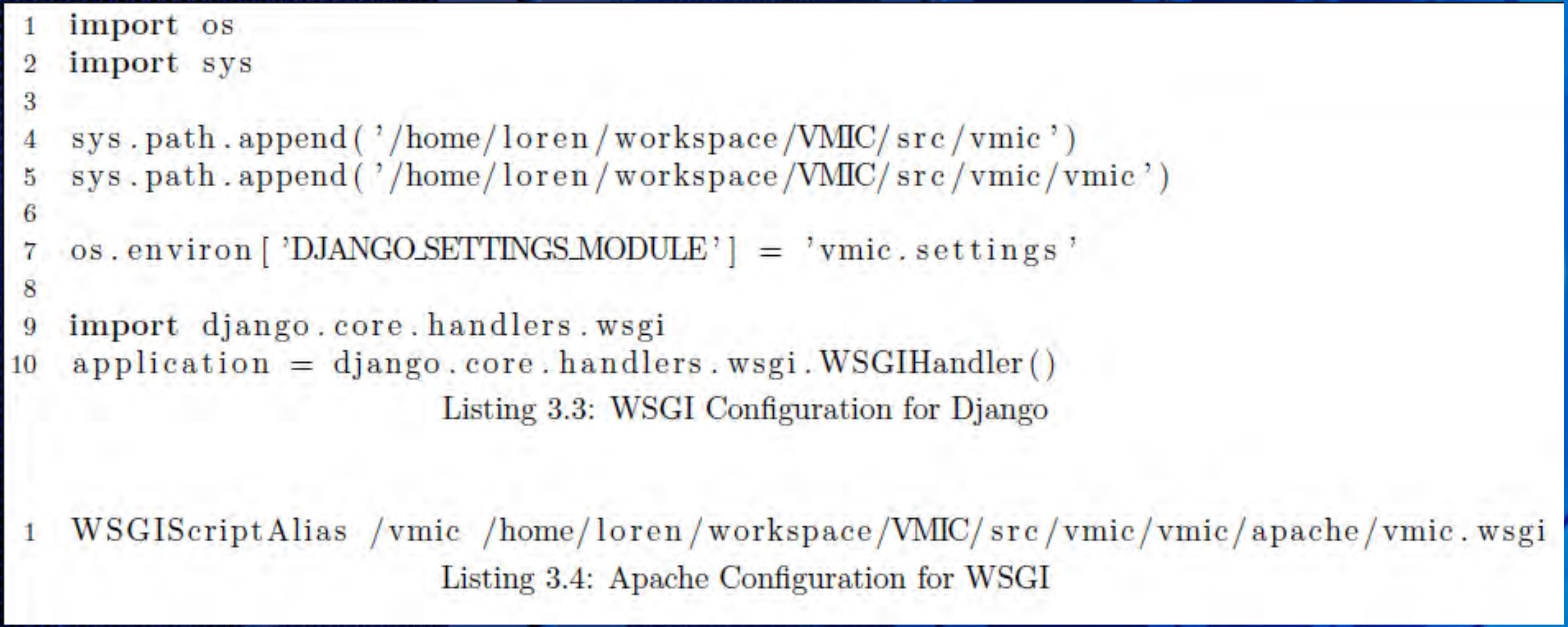
Recent Actions

My Actions

None available

# Technology - Django with Apache

- Must use WSGI (Web Server Gateway Interface)
- Developed by Python for Communication between Applications and Servers
- Fairly Simple with mod_wsgi and Apache



```
import os
import sys

sys.path.append('/home/loren/workspace/VMIC/src/vmic')
sys.path.append('/home/loren/workspace/VMIC/src/vmic/vmic')

os.environ['DJANGO_SETTINGS_MODULE'] = 'vmic.settings'

import django.core.handlers.wsgi
application = django.core.handlers.wsgi.WSGIHandler()
```

Listing 3.3: WSGI Configuration for Django

```
WSGIScriptAlias /vmic /home/loren/workspace/VMIC/src/vmic/vmic/apache/vmic.wsgi
```

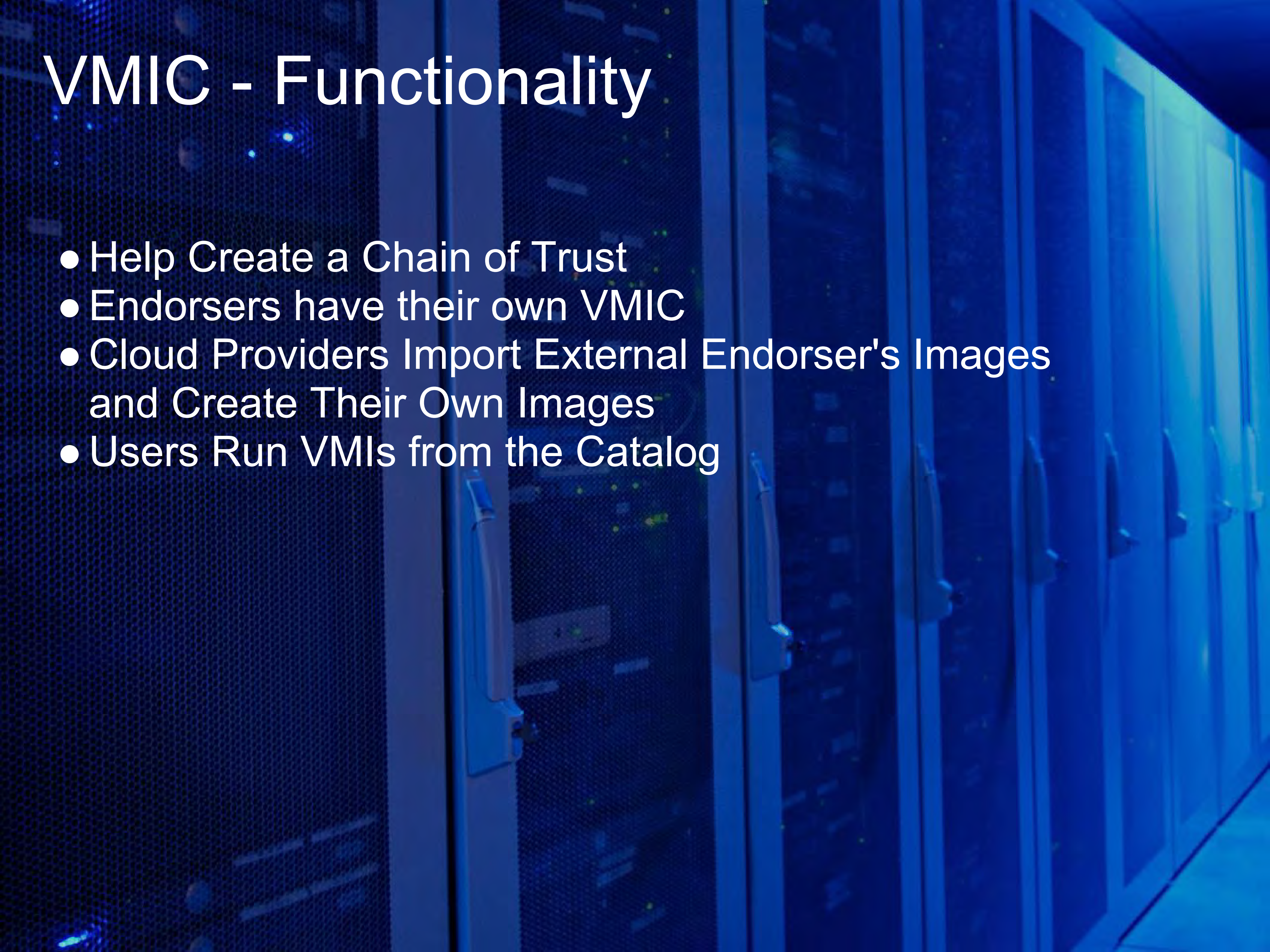
Listing 3.4: Apache Configuration for WSGI

# Technology - Django with Google AppEngine

- AppEngine provides automatic scalability
- No Servers to Worry About
- Requires an Account and SDK
- Requires a Nonrelational Database Backend (Not SQLite or MySQL)
- Can Require Code Workarounds if the Code uses Many to Many Relations

# Technology - Shibboleth

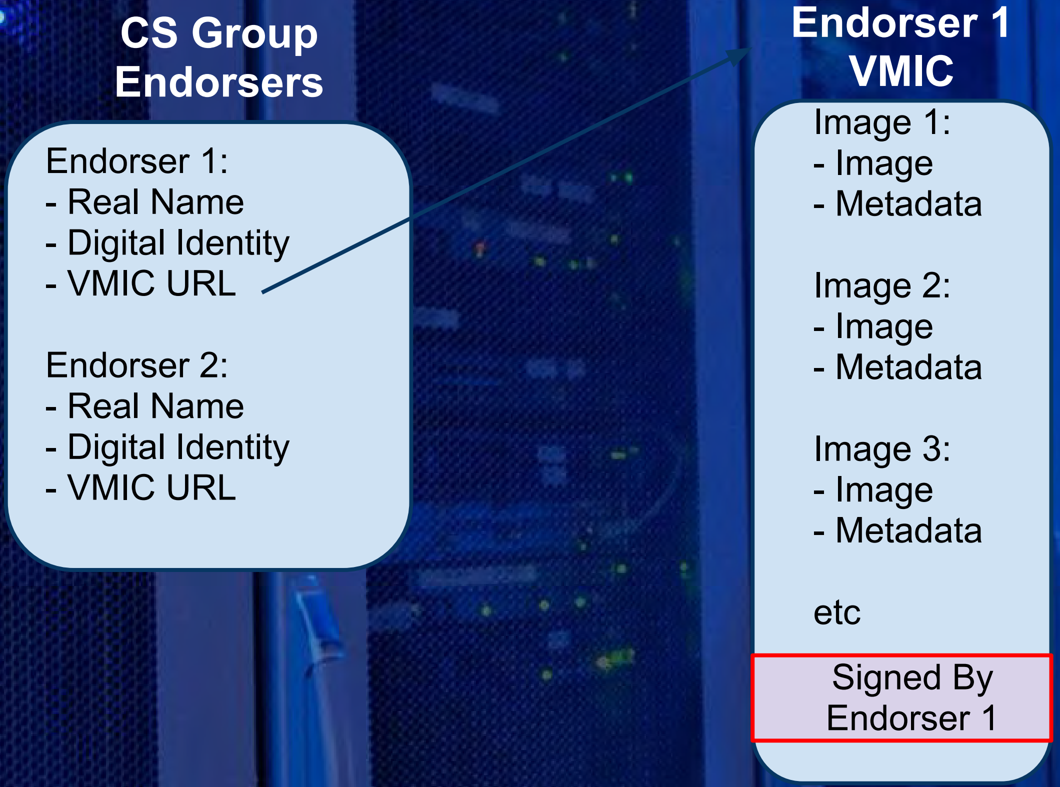
- Open Source, Standards Based System for Single Sign-on Across or Within Organizational Boundaries
- Used to allow Clemson Users to Sign on to VMIC
- Configurable to only Release Certain Membership Information (eg A Member of the University or Some Class, but not Name or any other Information)
- Client Systems Set Trusted Identity Providers
- Identity Providers Set Allowed Clients and Information to Release per Client
- mod_shib Integrates with Apache
- Secure Directories can be Set in .htaccess Files or Shibboleth Configuration Files



# VMIC - Functionality

- Help Create a Chain of Trust
- Endorsers have their own VMIC
- Cloud Providers Import External Endorser's Images and Create Their Own Images
- Users Run VMIs from the Catalog

# VMIC - Endorsers



**CS Group Endorsers**

Endorser 1:
- Real Name
- Digital Identity
- VMIC URL

Endorser 2:
- Real Name
- Digital Identity
- VMIC URL

**Endorser 1 VMIC**

Image 1:
- Image
- Metadata

Image 2:
- Image
- Metadata

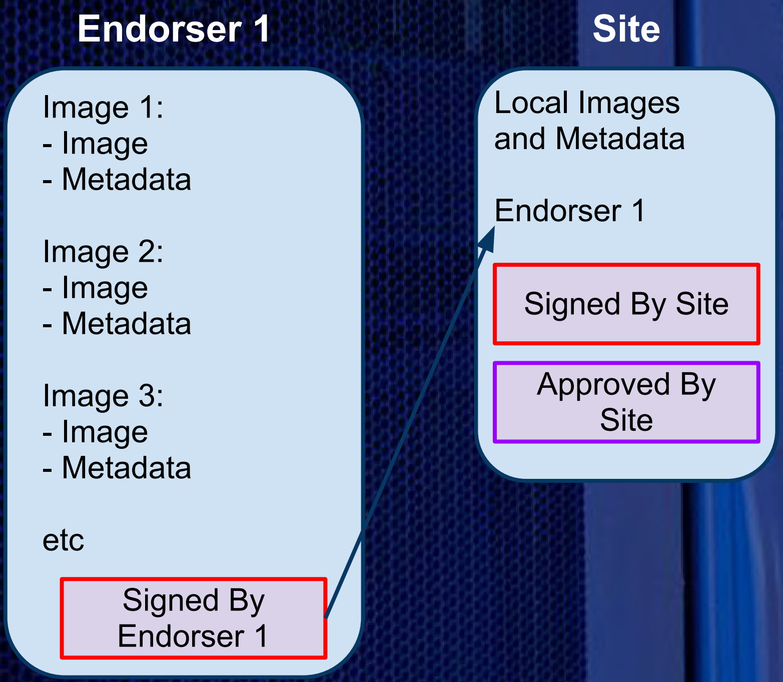
Image 3:
- Image
- Metadata

etc

Signed By Endorser 1

Each endorser publishes a VMIC which is signed with that endorsers digital identity.
The VMIC includes the endorsed images and some metadata about the images.
Endorsers say that the image meets a certain set of standards.

# VMIC - Getting Approved to Run at a Site



## Endorsed (endorser decision):

- Role defined in the policy document
- Scope: VMI production & maintenance

## Approved (site decision):

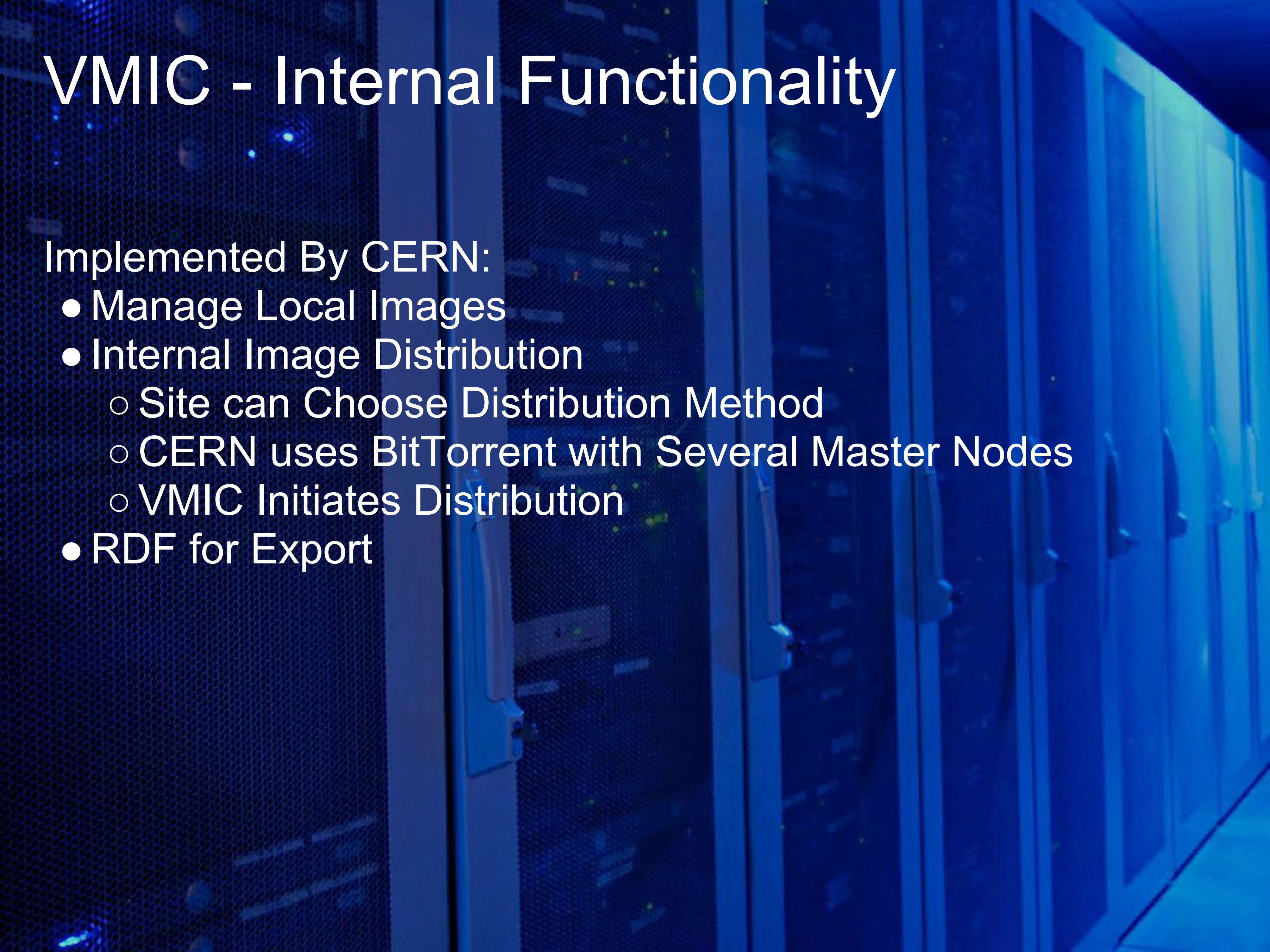
- Marks the VMI (or Endorser) “valid for use” by the site
- Scope: operating the VMI

## For a VMI to run, it must be both:

- Endorsed by an endorser
- Approved by the local site

The signatures indicate trust between the site and endorsers.

# VMIC - Internal Functionality



Implemented By CERN:

- Manage Local Images
- Internal Image Distribution
  - Site can Choose Distribution Method
  - CERN uses BitTorrent with Several Master Nodes
  - VMIC Initiates Distribution
- RDF for Export

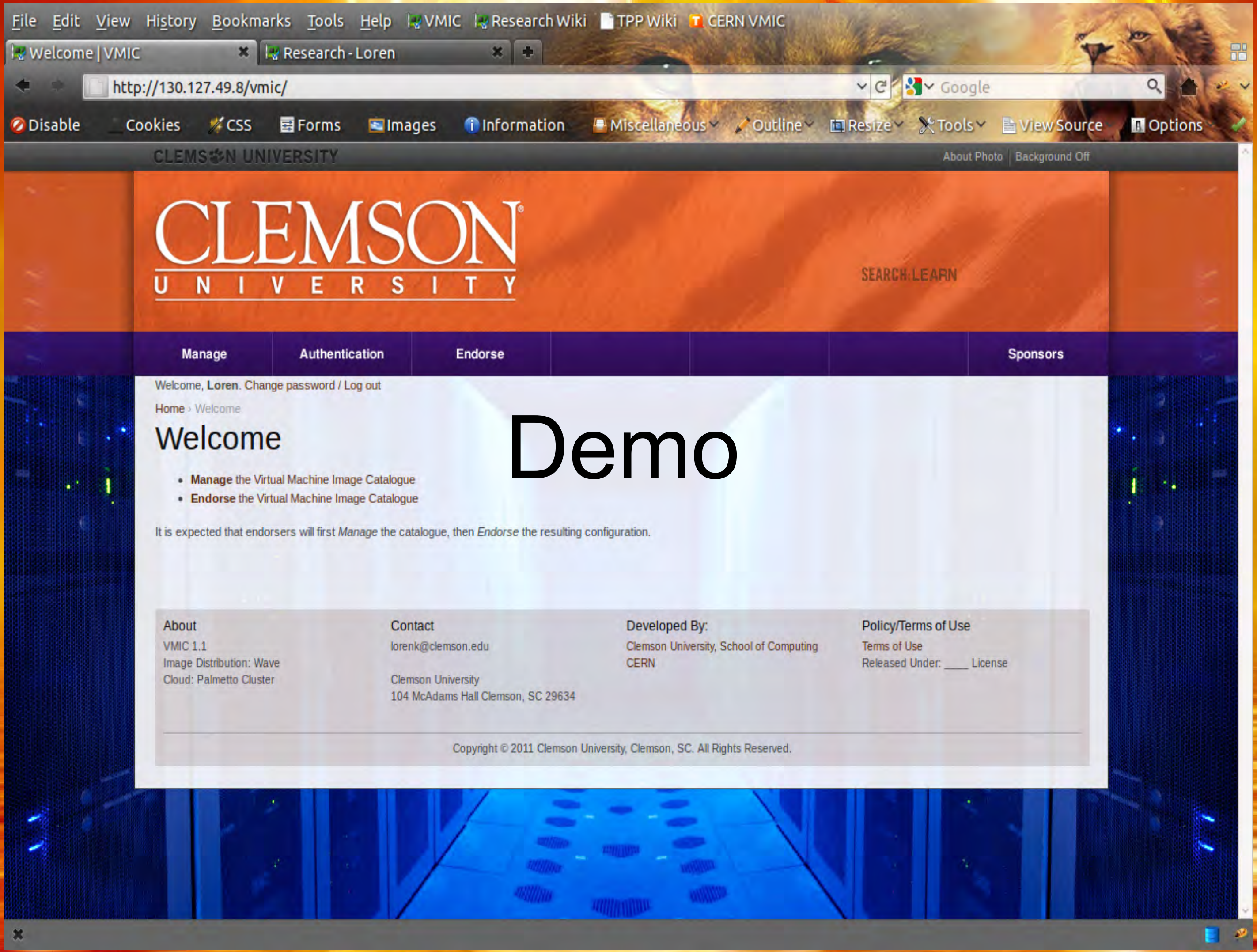
# VMIC - External Functionality Issues

- Image Validation
- Catalog Validation
- Image Updates from Endorsers
- VMI Export Chaining

# VMIC - External - Proposed Solutions

- Image Validation
  - Check File Hash (SHA-1, SHA-224, SHA-256, etc)
  - Don't Use MD5
  - SSL Transfer/Check
- Catalog Validation
  - Public Key Encryption
  - Public Key Infrastructure (Certification Authorities)
  - Certificate Validation
- Image Updates from Endorsers
  - Cron Job for Checking for Catalog Updates
  - Separate Job to Fetch New Images
- VMI Export Chaining
  - Not Allowed
  - Possibly Chain Endorsers in the Future

# VMIC - Prototype Implementation


File Edit View History Bookmarks Tools Help VMIC Research Wiki TPP Wiki CERN VMIC
Welcome | VMIC
Research - Loren
http://130.127.49.8/vmic/
Google
Disable Cookies CSS Forms Images Information Miscellaneous Outline Resize Tools View Source Options
CLEMSON UNIVERSITY
About Photo | Background Off
CLEMSON UNIVERSITY
SEARCH LEARN
Manage
Authentication
Endorse
Sponsors

Welcome, **Loren**. Change password / Log out

Home › Welcome

## Welcome

- **Manage** the Virtual Machine Image Catalogue
- **Endorse** the Virtual Machine Image Catalogue

It is expected that endorsers will first *Manage* the catalogue, then *Endorse* the resulting configuration.

About
VMIC 1.1
Image Distribution: Wave
Cloud: Palmetto Cluster

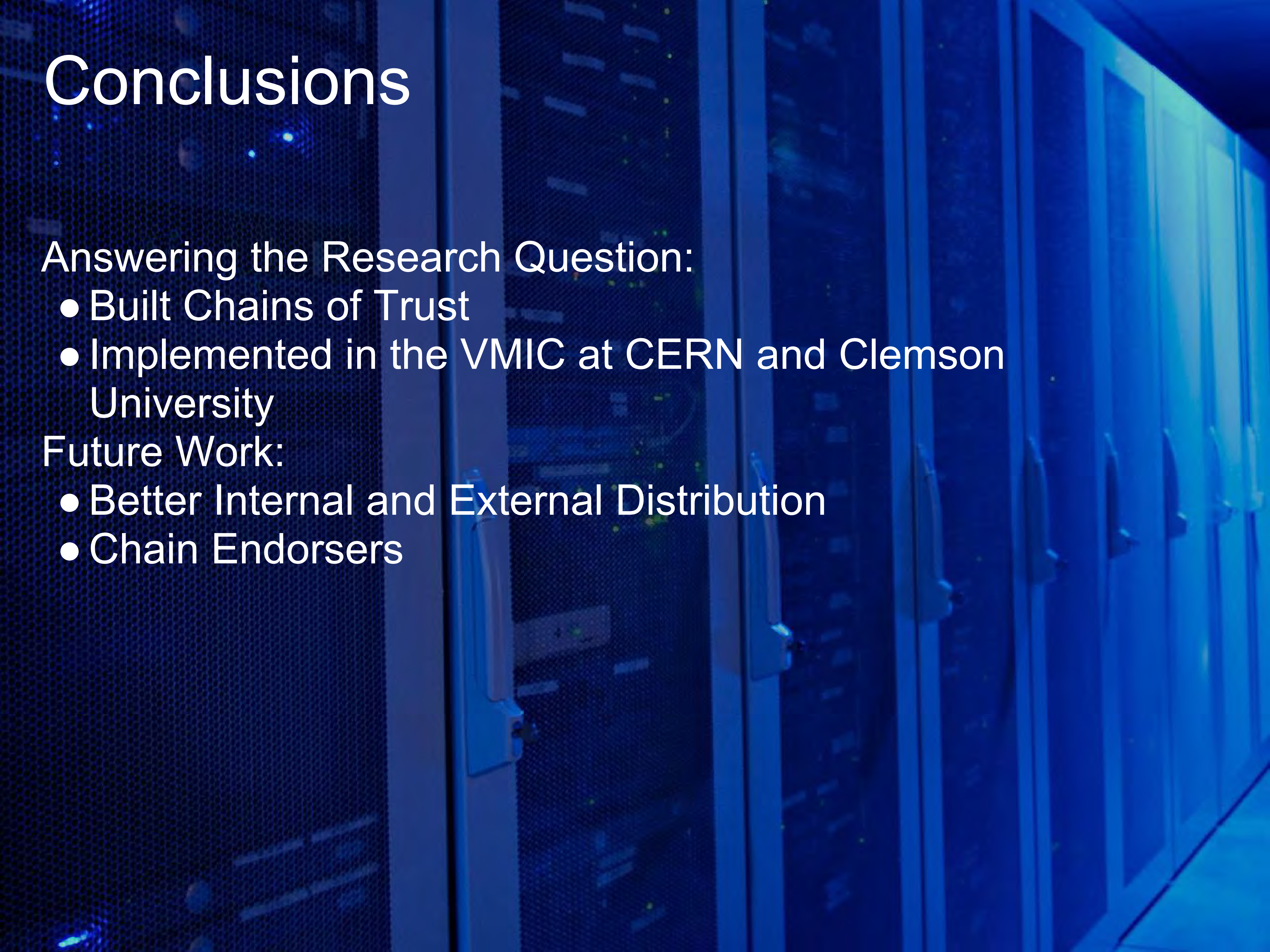
Contact
lorenk@clemson.edu

Clemson University
104 McAdams Hall Clemson, SC 29634

Developed By:
Clemson University, School of Computing
CERN

Policy/Terms of Use
Terms of Use
Released Under: ____ License

Demo



# Conclusions

Answering the Research Question:

- Built Chains of Trust
- Implemented in the VMIC at CERN and Clemson University

Future Work:

- Better Internal and External Distribution
- Chain Endorsers

# Acknowledgements

Dr. Goasguen - Research Advisor
CERN - Began VMIC
Honor's College - Some Funding

# Questions?