

A High Performance Virtualized Seismic Data Acquisition System

G.A. Davis¹, J.A. Eakins¹, J.C. Reyes¹, Mathias Franke², Rodrigo Sanchez³, Patricio Cortes Muñoz³, Robert W Busby⁴, Frank Vernon¹, Sergio Barrientos Parra³
¹Scripps Institution of Oceanography, La Jolla, CA, United States; ²Kinematics, Inc., Pasadena, CA, United States; ³Universidad de Chile, Centro Sismológico Nacional, Santiago, Chile; ⁴IRIS, Washington, DC, United States

INTRODUCTION

The **Centro Sismológico Nacional (CSN)** at the University of Chile is tasked with 24-hour monitoring of seismic stations throughout the country, making it available to authorities, particularly though agencies responsible for improving building standards.

Using lessons learned from the USArray Transportable Array, representatives from IRIS and Kinematics assisted the CSN in designing, installing and configuring a virtualized seismic data acquisition and processing system. The main installation was performed during July of 2014. Data from existing sensors as well as newly deployed instruments is acquired and processed by this system.

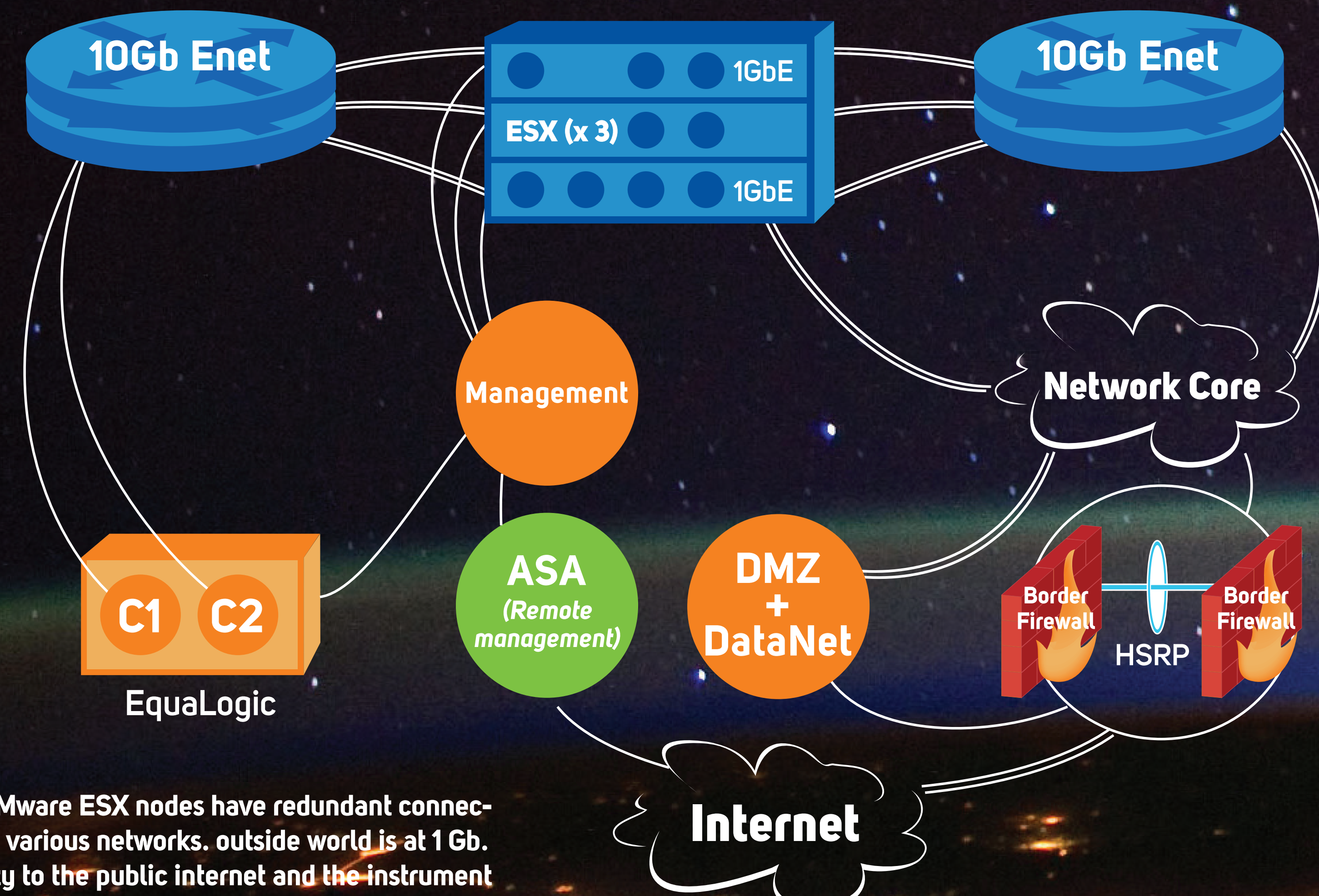
DESIGN GOALS

The system should provide protection against physical server and network path failure.

To guard against potential processing bottlenecks affecting timely data acquisition and exchange, the acquisition, processing, and data exchanges components should run in insulated containers.

The hardware and virtual systems should be remotely administrable via a dedicated management VPN, separate from the data path.

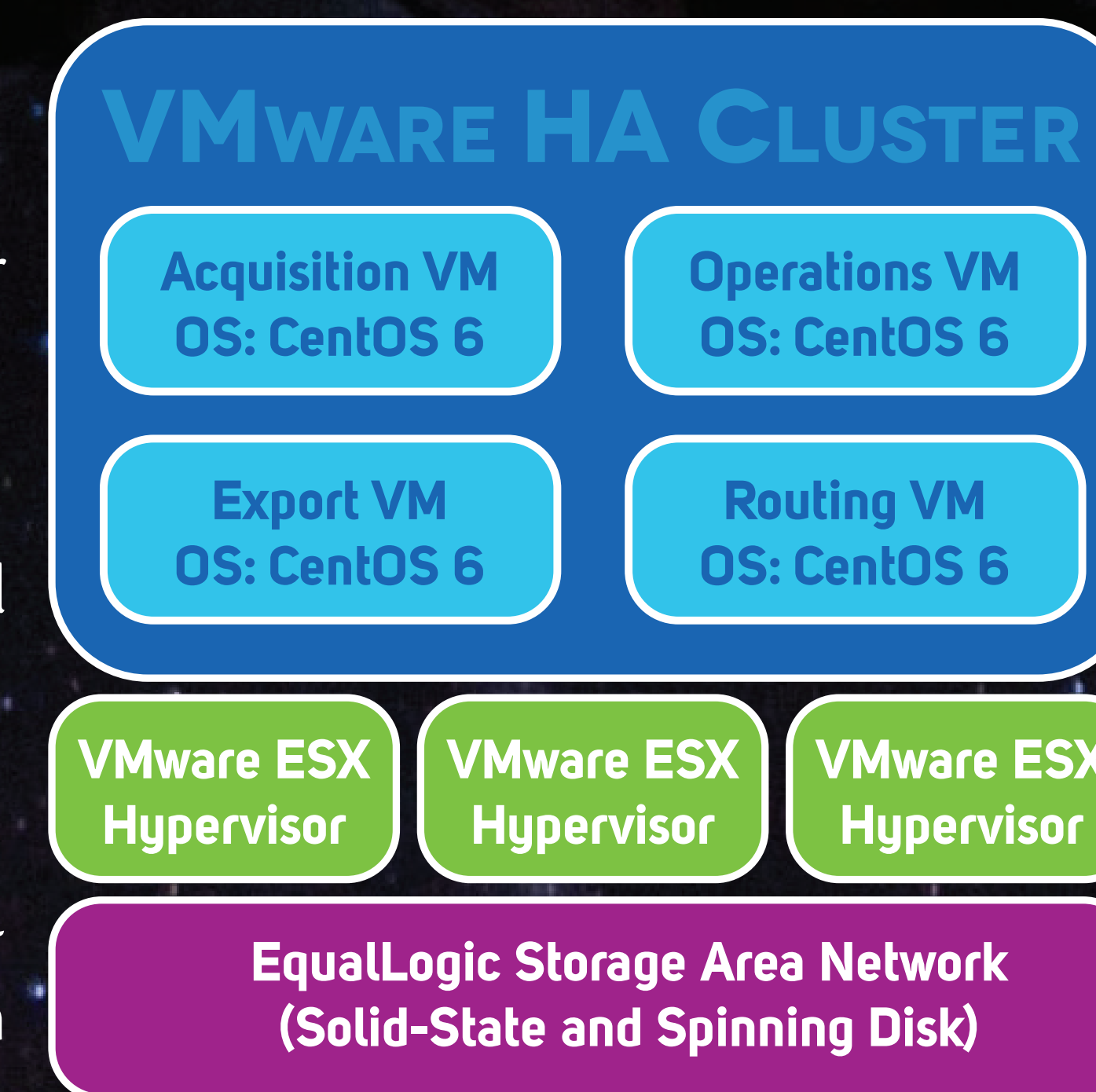
Finally, the systems should exchange data with existing data processing systems in a timely fashion.



All three VMware ESX nodes have redundant connections to the various networks. outside world is at 1 Gb. Connectivity to the public internet and the instrument network is through the existing network core.

SERVERS AND INFRASTRUCTURE

- 3x Dell R720 servers
- 4x Force10 10GigE network switches
- GigE network for management LAN, separate from server networks
- Cisco ASA Firewall for management LAN access
- EqualLogic Storage Area Network (SAN) - 20 TB Hybrid Solid State and Spinning Disk with automatic data tiering.
- VMware vSphere 5.5 in HA cluster mode



Block diagram showing the major components of the VMware stack. The VMware images and the large data volume are located on the shared SAN storage.

All hardware installed for this project integrates with the existing network at CSN. Each of the network segments are built with redundant physical paths all the way through the network core out to the perimeter firewalls.

OPERATING ENVIRONMENT

The core operating environment consists of three CentOS 6.5 Virtual Machines running BRTT, Antelope 5.3. In order to handle some of the more esoteric routing requirements to stations on private VSAT links or restricted access networks, a single (highly-available) pfSense virtual machine was included in the design.

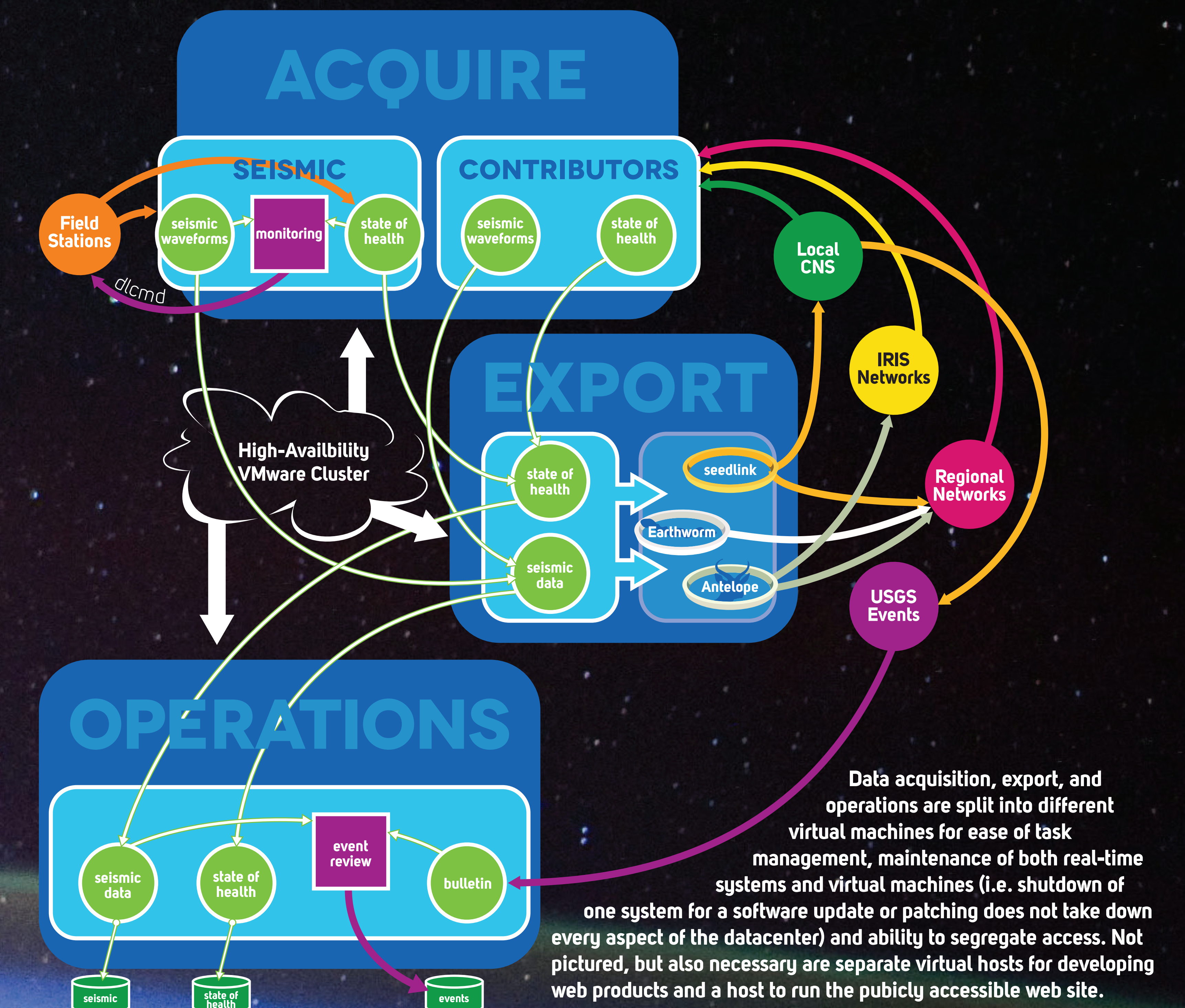
There are two volumes provisioned on the SAN: a 5 TB volume formatted as VMFS which contains the VM images and their local processing information, and a 15 TB volume for data archiving. The archive volume is configured as a pass-through LUN and is mounted directly on the operations system. The data is shared via NFS to the acquisition and export nodes, as well as to some of the legacy systems at CSN. The pass-through option was chosen in order to allow the volume to be hosted by a dedicated NFS server at a later date.

DATALOGGERS

The Antelope acquisition system connects to 30 Kinematics Basalt dataloggers and 10 Quanterra Q330 digitizers, managing them, and acquiring seismic data as well as State-of-Health (SOH) information. Stations from CSN's legacy network and contributing networks are acquired from a separate real-time system on the acquisition node.

SOFTWARE CONFIGURATION

Acquisition, export, and processing systems are split onto different CentOS 6 virtual machines, and into different functional groups on those VMs. Data flows in from stations in the field and contributed networks through the acquisition node, and is then distributed to the operations system and legacy processing infrastructure through the export node. Databases are written by the operations host, which also provides access via NFS to the other nodes.



Data acquisition, export, and operations are split into different virtual machines for ease of task management, maintenance of both real-time systems and virtual machines (i.e. shutdown of one system for a software update or patching does not take down every aspect of the datacenter) and ability to segregate access. Not pictured, but also necessary are separate virtual hosts for developing web products and a host to run the publicly accessible web site.