A Desktop Grid Network design for HSI applications

Graduated Student: José O. Nogueras Colón

Adviser: Yahya M. Masalmah, Ph.D.

Agenda

Introduction

- Problem Statement
  - Hyperspectral imagery (HSI) applications involve large data storage and computational power on real time. Desktop Grid Computing has become a powerful solution for the current computational challenges. On the other hand, Cloud Computing offers high storage scalability into a virtualization environment. Our research outlines a review of the best practice existing for desktop grid and cloud computing. We propose a secure network design for HSI using the best of desktop grid and cloud infrastructure.

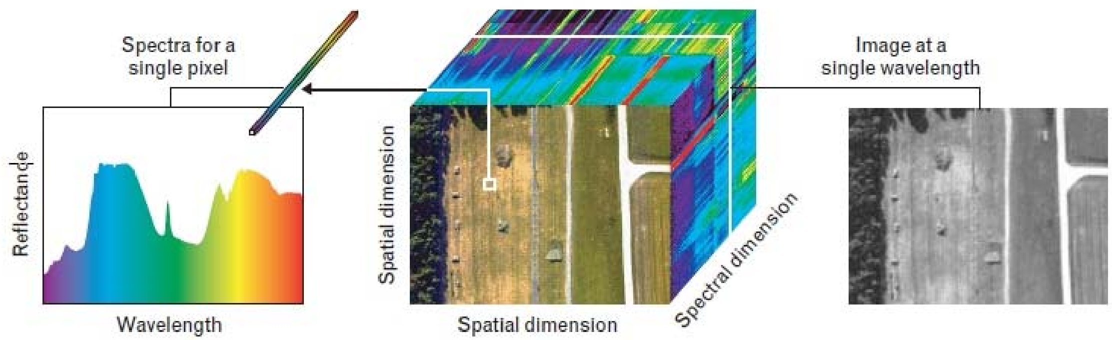
Objectives

- Network design
  - Implementation
  - Assessment and Validation
- Porting our HSI application, Contrained Positive Matrix Factorization (cPMF) algorithm, on a Desktop Grid (DG) environment.
- Application deployment and performance analysis in comparison with the most recent methods used.

Hyperspectral Imagery (HSI)

- What is a Hyperspectral Image?

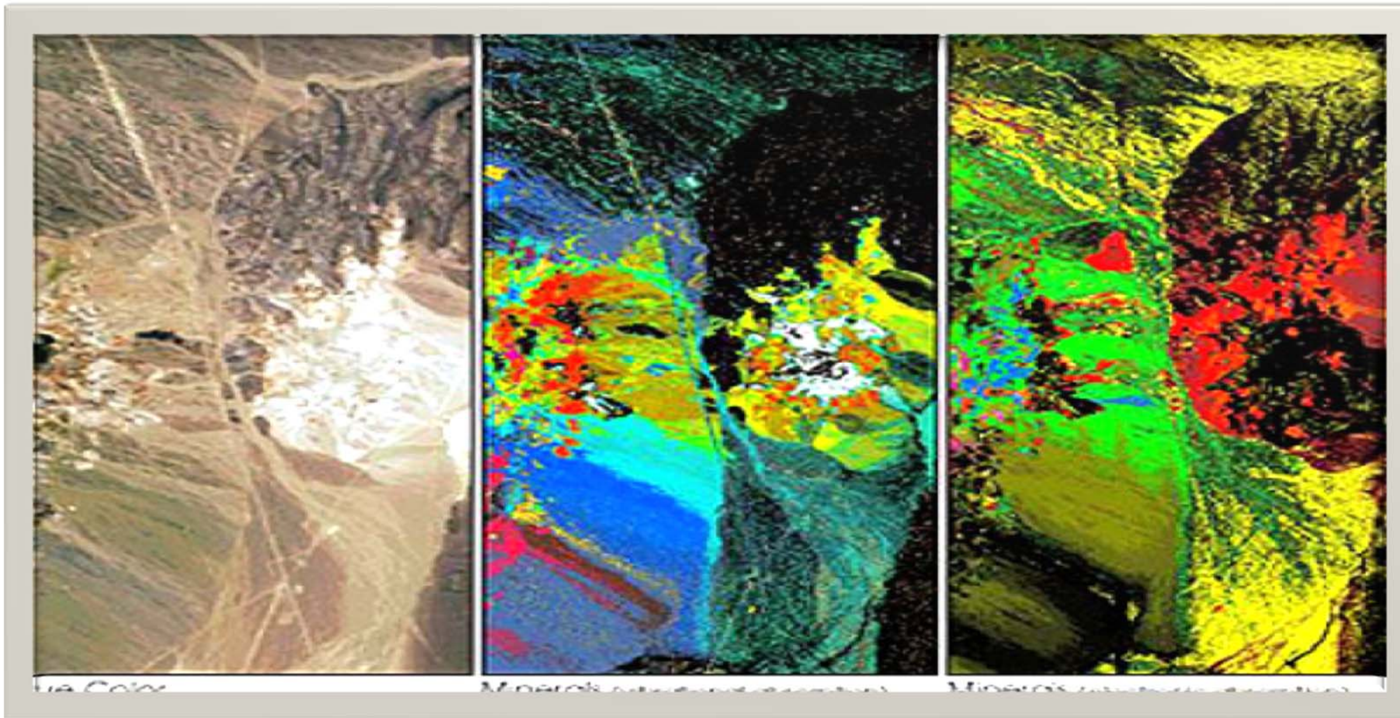
Hyperspectral sensors collect data across a wide range of the spectrum (VNIR–LWIR, plus TIR) at small spectral resolution (5–15 nm) and high spatial resolution (1–5 m). This allows detailed spectral signatures to be identified for different imaged materials. HSI describe the imagery derived from subdividing the electromagnetic spectrum into very narrow bandwidths. These narrow bandwidths may be combined with or subtracted from each other in various ways to form images useful in precise terrain or target analysis.



Applications

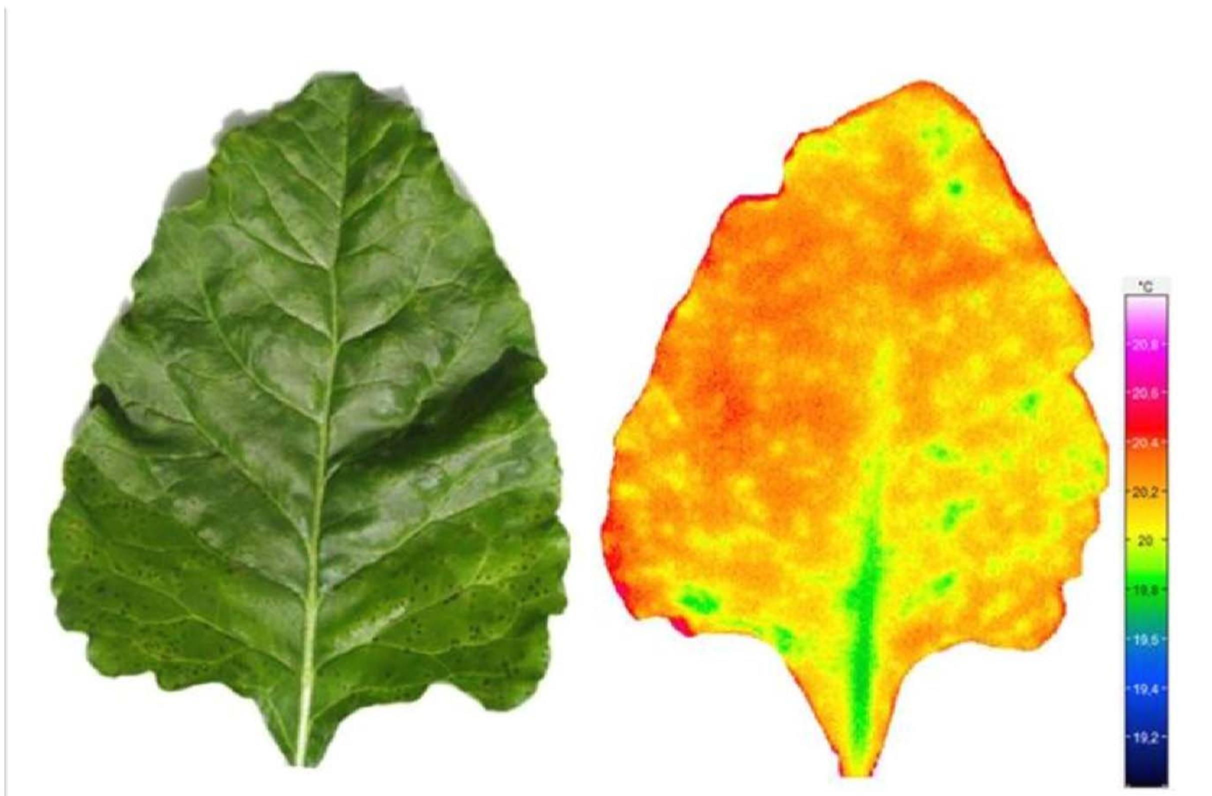
Examples of HSI Applications

- Mineralogy
  - Used to detect what minerals are present in an image



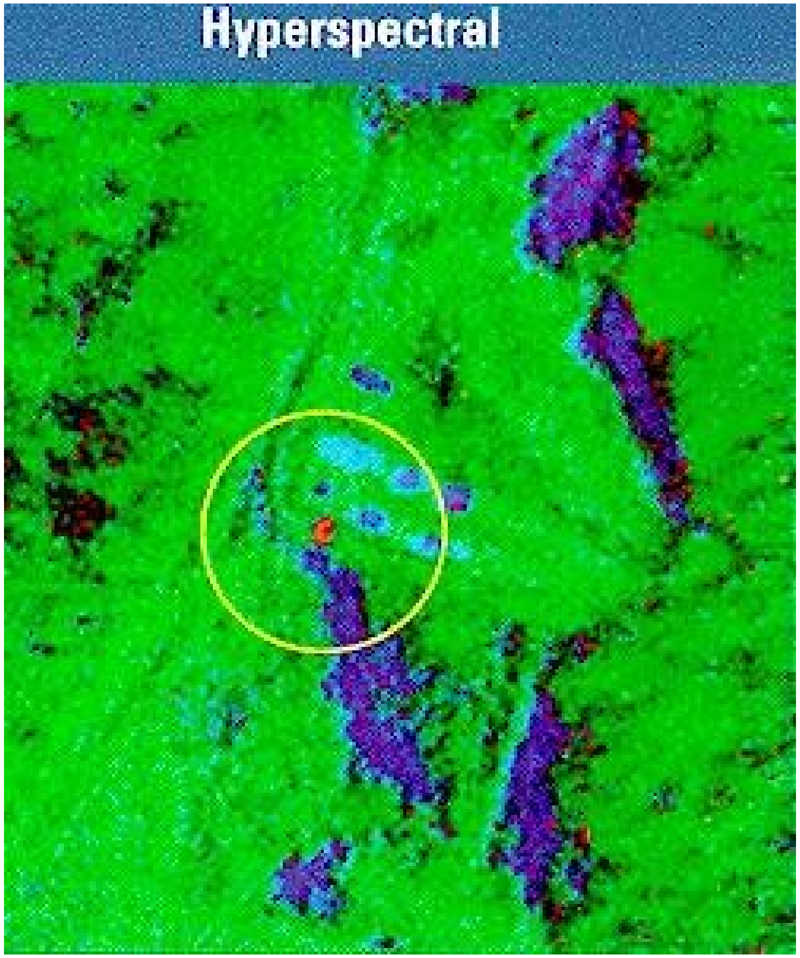
Examples of HSI Applications

- Agriculture
  - Used to monitor crops health



Examples of HSI Applications

- Military Surveillance
  - Used to detect targets that are well hidden



Background

- Unmixing
  - Many pixels in the scene contain several sub-pixel level, thus resulting spectral signature in these pixels is not "pure" but is given by a composition or mixture of different substances that coexist at the sub-pixel.

- The excessive computational process.

Our Application for Target Detection: cPMF

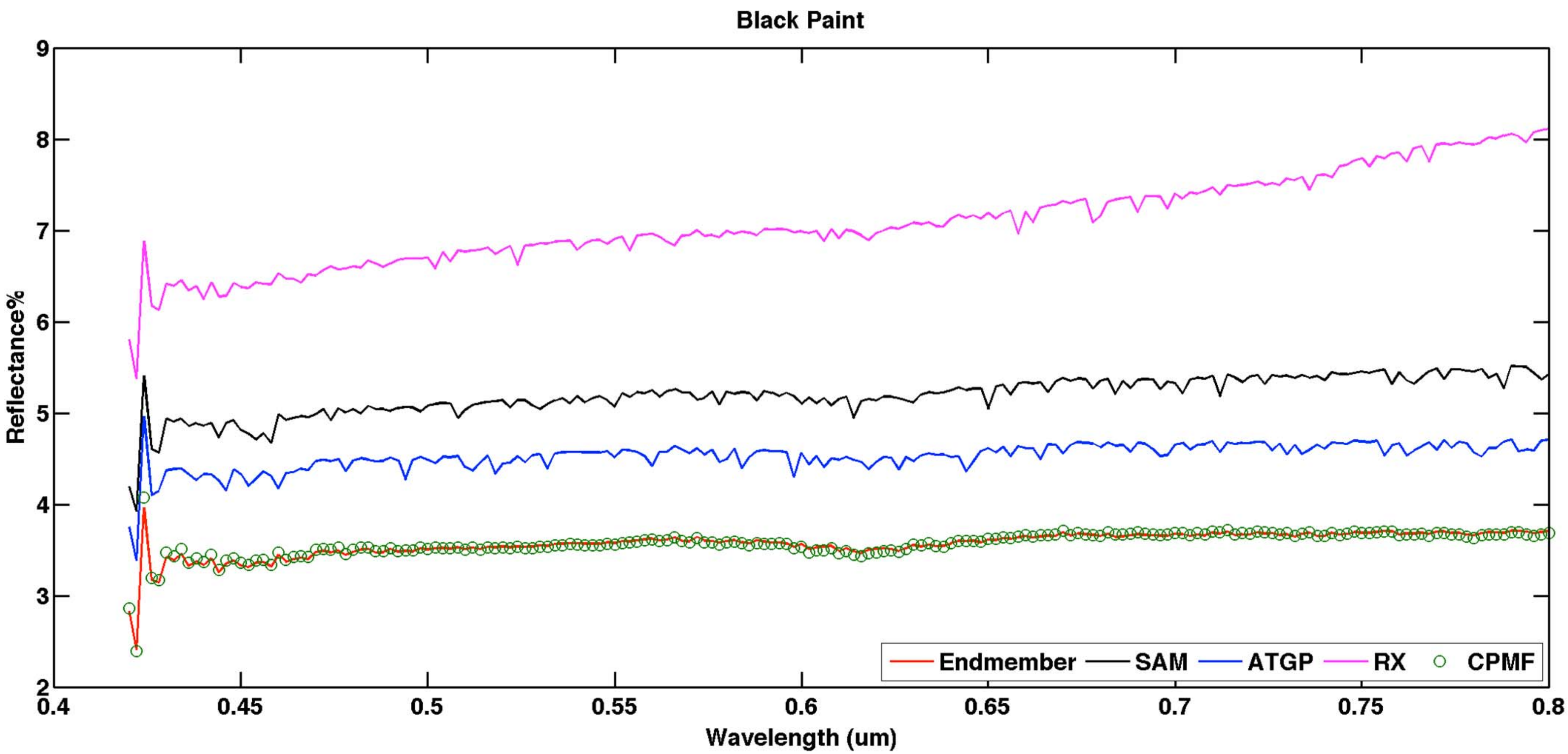
- The Constrained Positive Matrix Factorization algorithm (CPMF) pretend solve the unsupervised unmixing problem by extracting simultaneously, the endmembers, and their abundances.
- CPMF consists of three different steps: determination of number of endmembers, initialization, and computing the constrained positive matrix factorization.
- This algorithm does not assume that endmembers are pure pixels in the image.

Most recent results

- Speed

Algorithm	SAM	RX	ATGP	CPMF
Time Elapsed (sec)	1.4000	2.3300	2.6600	4806

Cont…

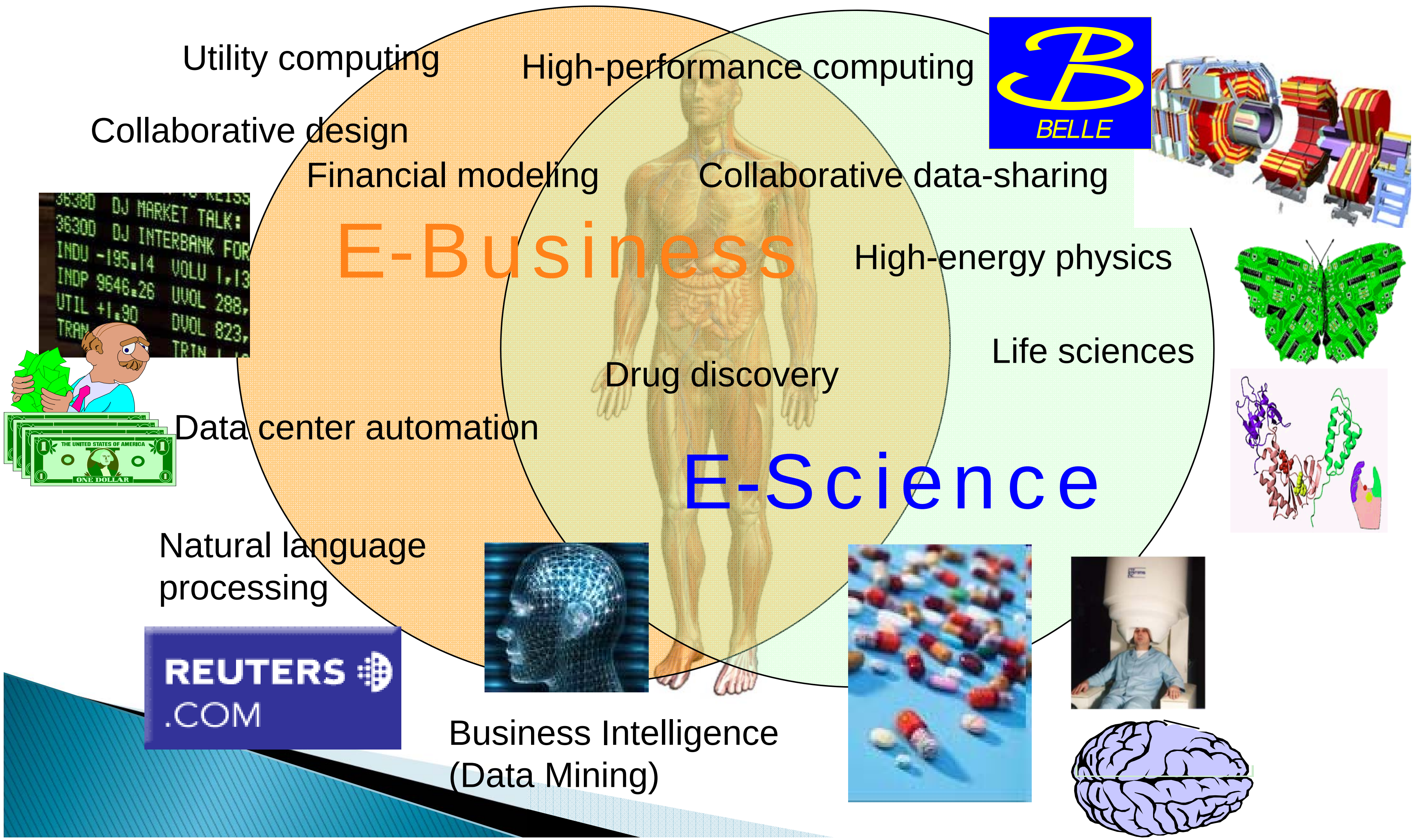


Black Paint
Reflectance%
Wavelength (um)
9
8
7
6
5
4
3
2
0.4
0.45
0.5
0.55
0.6
0.65
0.7
0.75
0.8
Endmember
SAM
ATGP
RX
CPMF

Grid Computing

- Allows sharing and coordinated use of diverse resources of many computers in a network to a single problem at the same time – usually to a scientific or technical problem that requires a great number of computer processing cycles or access to large amounts of data [7].
- These resources may include processing power, storage capacity, specific data, and other hardware such as input and output devices.

How Are Grids Used?

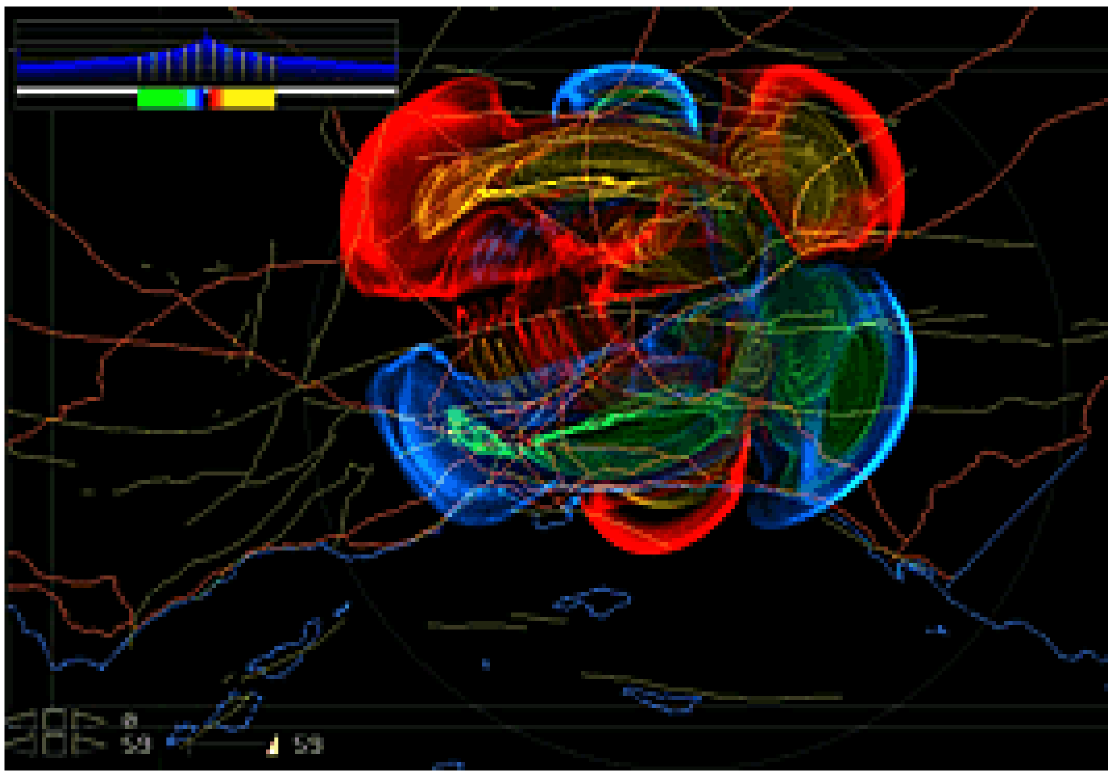


Types of Grid Computing

- The first type are heavy-weight, feature-rich systems that tend to concern themselves primarily with providing access to large-scale, intra- and inter-institutional resources such as clusters or multiprocessors.
- The second general class of Grid computing systems is the Desktop Grids, in which cycles are scavenged from idle desktop computers [7].

Applications

- The *Southern California Earthquake Center* uses Globus software to visualize earthquake simulation data.



Desktop Grids

- In a desktop grid system, the execution of an application is orchestrated by a central scheduler node, which distributes the tasks amongst the worker nodes and awaits workers' results [1][8].

- Desktop grids fall into two categories:
  - Local
  - Volunteer

DGs Advantages

- Heterogeneous systems
  - Hardware and Software
- Low cost
  - Desktops vs. Supercomputers and complex clusters systems cost.
  - Less Energy Consumption
  - Open Source Middleware and Frameworks
- Allow anyone to join
  - Flexible Scalability
- Great computational power for e-science
- Green IT
- Easy Maintenance and Support

Green DGs

- The Green key advantage of Desktop Grids over service Grids and data centers based on clusters of servers is the minimal heat density.

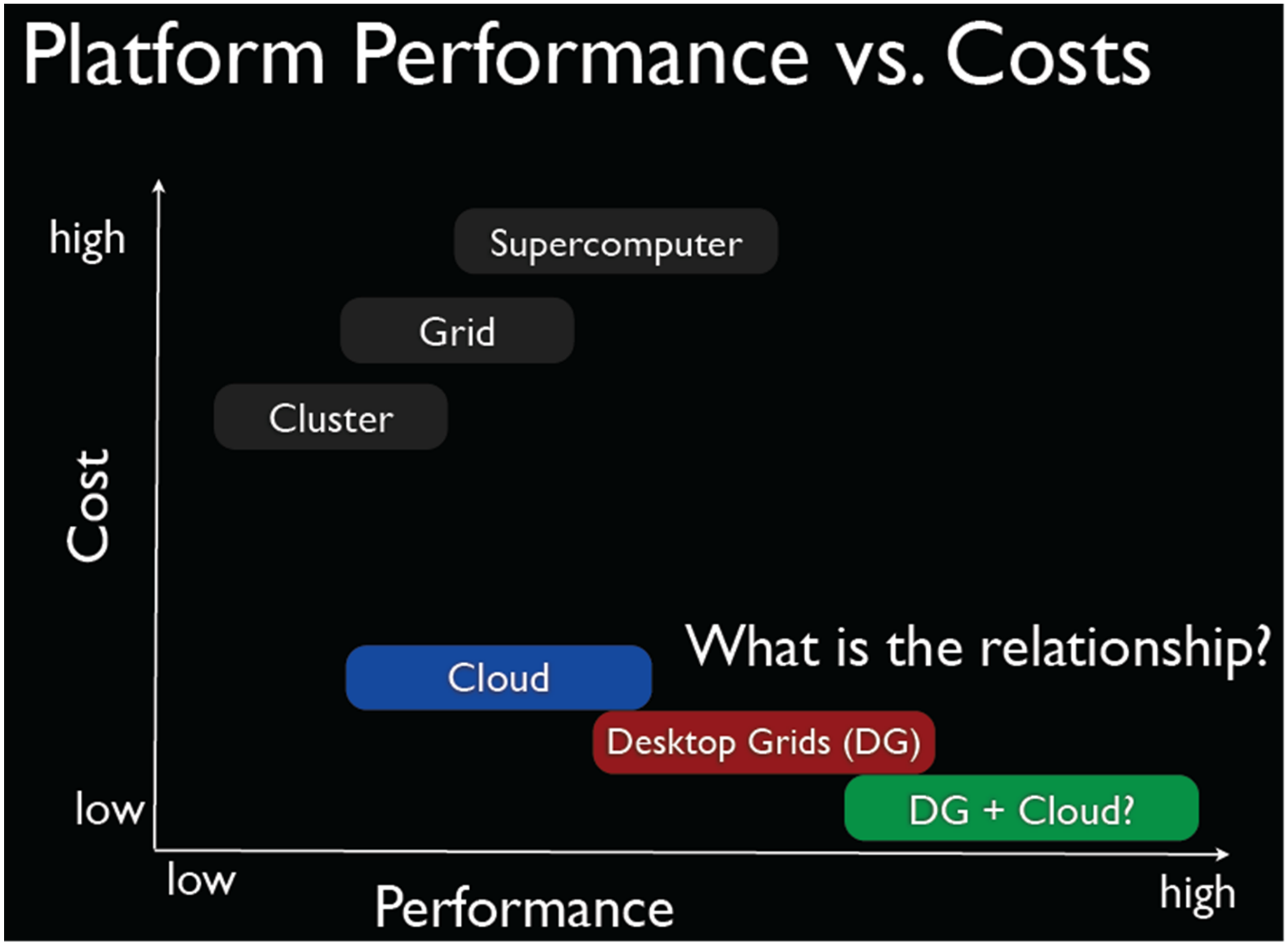
- Compute Clusters without energy intensive air condition would run into thermal disaster within minutes. PCs participating in Desktop Grids usually do not make use of any air condition.

- Schott and Emmen (2011), investigates whether Desktop Grids still have a Green advantage over Service-Grids and describe several distinct Green Methodologies to optimize compute unit specific energy consumption. Green-IT metrics as Carbon Footprint and power usage efficiency (PUE) were analyzed for their relevance and applied to Desktop-Grids [9].

Desktop Grids + Cloud Computing

- *Cloud computing* offers “infinite” resources on demand [2].
- Cloud Computing offers high storage scalability into a virtualization environment.
- Benefits:
  - Pay as you go.
  - Scale up or down dynamically.
  - No hardware management, less software management.

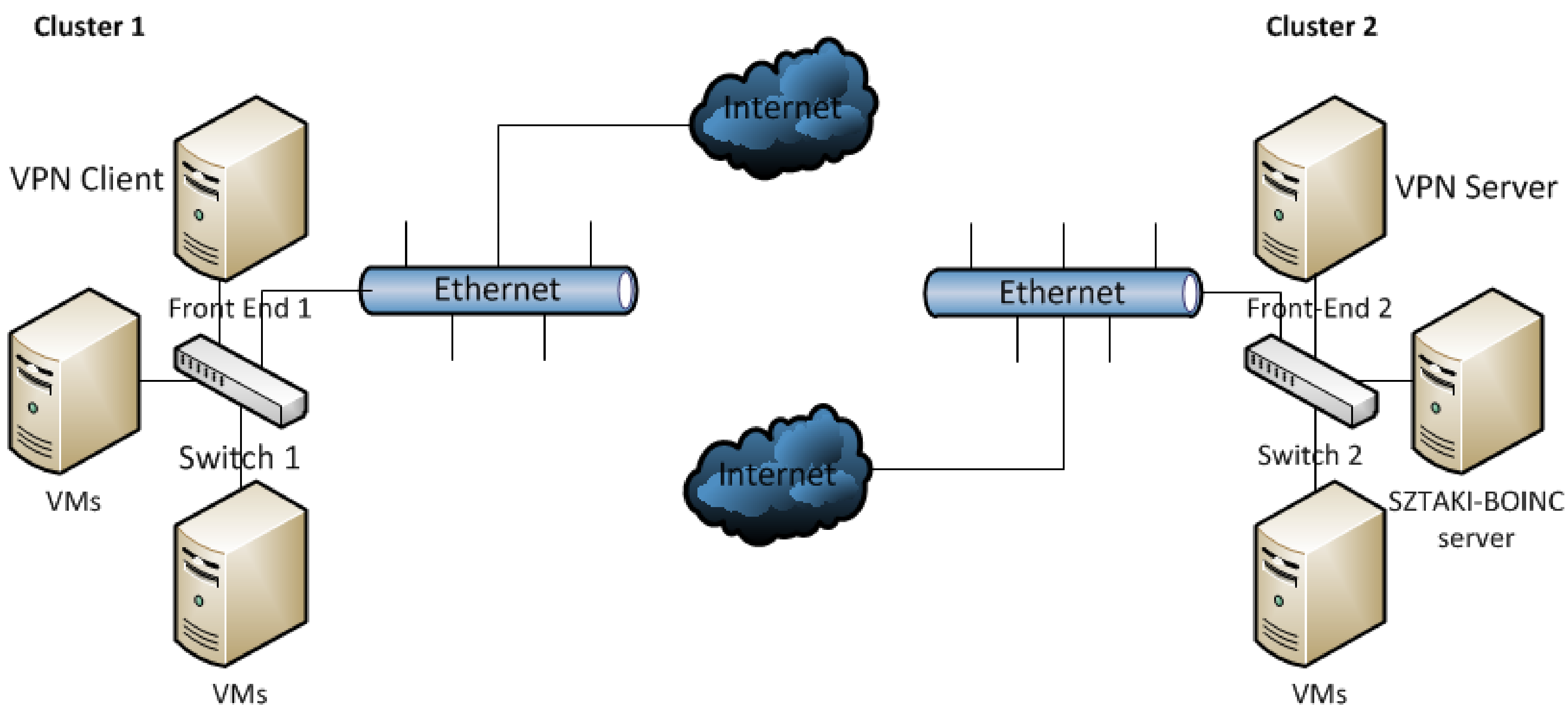
Cont...


Platform Performance vs. Costs
high
Supercomputer
Grid
Cluster
Cost
What is the relationship?
Cloud
Desktop Grids (DG)
low
DG + Cloud?
low
Performance
high

DGs Network design Considerations

- Bandwidth
- End-to-End Latency
- Maximun Transfer Unit Size
- Data Management
- Security issues

Network Design



Grid on Cloud computing
Physical Network Design

Conclusion

- Discussed about what is HSI, its applications and our target detection algorithm cPMF.
- We did a review of the best practice methods on Desktop Grid Computing and its advantages.
- A DGs physical network design has been proposed.

Future Work

- Improve the CPMF speed (sequencial and distributed)
- Compares the currently CPMF language (Matlab) with others programming languages like C++, Phyton and Cuda.
- Assessment and Validation of our network design: Simulated and Real.
- DG remote access for Volunteer Computing.

Acknowledgement

- This research is supported by Department of Energy under Samuel P. Massie Chair of Excellence: Sub-Project #3: “Development of Serial and Parallel Signature and Material Extraction Algorithms from Advanced HSI Data Sets”

References

- [1] Z. Balaton, G. Gombás, P. Kacsuk, A. Kornafeld, J. Kovács, A. C. Marosi, G. Vida, N. Podhorszki and T. Kiss. SZTAKI desktop grid: A modular and scalable way of building large computing grids.
- [2] M. B. Belgacem, H. Hafsi and N. Abdennadher. "A hybrid Grid/Cloud distributed platform: A case study," in *Grid and Pervasive Computing*Anonymous 2013, .
- [3] T. Ercan. Effective use of cloud computing in educational institutions. *Procedia-Social and Behavioral Sciences 2(2),* pp. 938-942. 2010.
- [4] T. Kiss, I. Kelley and P. Kacsuk. Porting computation and data intensive applications to distributed computing infrastructures incorporating desktop grids. Presented at The International Symposium on Grids and Clouds and the Open Grid Forum, PoS (ISGC 2011 & OGF 31). 2011, .
- [5] D. Kondo, B. Javadi, P. Malecot, F. Cappello and D. P. Anderson. Cost-benefit analysis of cloud computing versus desktop grids. Presented at Parallel & Distributed Processing, 2009. IPDPS 2009. IEEE International Symposium on. 2009, .
- [6] E. J. Korpela. SETI@ home, BOINC, and volunteer distributed computing. *Annu. Rev. Earth Planet. Sci. 40*pp. 69-87. 2012.
- [7] R. Kumar, I. Khan and V. Gupta. Literature review on grid computing. *African Journal of Mathematics 6(7),* pp. 144-148. 2013.
- [8] A. Marosi, Z. Balaton, P. Kacsuk and D. Drotos. SZTAKI desktop grid: Adapting clusters for desktop grids. *Remote Instrumentation and Virtual Laboratories* pp. 133-144. 2008.
- [9] B. Schott and A. Emmen. Green desktop-grids: Scientific impact, carbon footprint, power usage efficiency. *Scalable Computing: Practice and Experience 12(2),* 2011.
- [10] B. Schott and A. Emmen. Green methodologies in desktop-grid. Presented at Computer Science and Information Technology (IMCSIT), Proceedings of the 2010 International Multiconference on. 2010, .