

# Enterprise Grid Requirements

**Jikku Venkat  
CTO  
United Devices  
GGF 12 Workshop  
21-Sep-2004**

Recipient of  
ComputerWorld's 2004  
21st Century  
Achievement Award  
*"Judged best IT  
application in the world  
in Medicine."*



# Contents

- **Background**
- **Sample Deployments**
- **Requirements Summary**
- **Conclusions**

# Our Vision of the Enterprise



## INFRASTRUCTURE TODAY

*Generally accepted that the average node's utilization is 5-10%, yet enterprises are constantly buying more infrastructure*

## INFRASTRUCTURE WITH UD

*Grid MP virtualizes access to infrastructure based on business priorities and policies with guaranteed service levels, generates huge ROIC while accelerating business processes*

## VALUE

*"We invested roughly \$400K in Grid software and figure we saved at least \$2 million" – Manuel Peitsch, Novartis*

## ENTERPRISE OPTIMIZATION

*Grid MP enables organizations to deploy an Enterprise-wide Grid across ALL compute infrastructure enabling IT professionals to manage and direct those resources for the first time*

# Company

- **History**

- Founded in 1999, headquarters in Austin, Texas with offices in Pennsylvania, New Jersey, Minnesota, UK and India
- Leading team from Dell, IBM, Microsoft, Intel, Seti@home and distributed.net

- **What we do**

- Develop and sell software and solutions that enable enterprise-wide grids – Grid MP Enterprise
  - Grid MP Workstation
  - Grid MP Cluster
  - Grid MP Data Center
  - MP Insight

- **Value Delivered**

- UD is a leader in creating dynamic compute infrastructures for research, development and business applications
  - Select Customers: Novartis, GSK, Johnson & Johnson, Sanofi-Aventis, DoD, Rapid Prototypes, Purdue University, University of Florida, P&G, Cinergy, MITRE...
  - Application Partners: Accelrys, Tripos, Schrodinger, syncsort, Landmark, LSTC,...
  - OEM Partners: IBM, HP, Sumisho Electronics, NTT Data, Singapore Computer Systems, Satyam, Infosys, .....

# Sample Deployments

# Sanofi~Aventis

- **Goal**
  - Screen entire 1M compound library
  - Invest in latest technology – skip over Clusters to Grids
- **Challenges**
  - Too costly to expand existing HPC systems (SGI)
  - Plans in place to phase out HPC systems across company
- **Solution Highlights**
  - Joint effort in partnership with Accelrys, Tripos to deliver integrated virtual screening solutions
  - Within 2 months began routinely screening 300k compounds across multiple locations in France
  - Recently completed a million compound screen
- **Future Plans**
  - Plan to expand to several thousand CPUs
  - Integrate other applications in discovery

sanofi~synthelabo

“Ease of deployment and the availability of applications were strong selling points for us. We also needed proven scalability and security and knew from United Devices’ other enterprise deployments and their work with their public Grid that massive scaling and security capabilities were already proven.”

Olivier Gien  
Head of Discovery IT  
Sanofi-Synthelabo

# GlaxoSmithKline

- **Goal**

- Replace internally developed Grid technology with commercially available solution

- **Challenges**

- Very knowledgeable about Grid computing having built VCS in-house
- Existing Platform Computing customer
- Very concerned with integration in their current IT infrastructure, application migration and standards.

- **Solution Highlights**

- Hosted Pilot allowed GSK to try UD solution with very little investment
- Rapid initial migration of existing applications
- No performance issues from non-dedicated nodes
- Running several applications (~15) in production on a few thousand nodes
- Applications in discovery and in clinical development

- **Future Plans**

- Plans to expand throughout enterprise including servers
- UD being certified as part of server image
- Investigating grid-enabled data warehouse solution

# *The Grid Driven Enterprise is Happening Now*

"If you look at dedicated clusters versus a more flexible [grid] environment where you can take advantage of both new acquisitions and existing systems through a single tool, it's just a better way to invest your money." J&J's grid technology of choice is the Grid MP platform from United Devices Inc., Austin, Texas. **"By going with the United Devices tool, we can have a single tool that can both do a virtual cluster and also take advantage of CPU harvesting off of the existing equipment that we already have.** That was one of the big reasons we picked United Devices over other ways of providing HPC capabilities."

"[Our grid] is a mixture of devices. Back-end servers, NPI devices—It's a way of taking a set of resources and treating them as if they were a cluster... **The United Devices tool gives us the ability to have all of these resources in the grid and through profiling and policy settings use them in different ways.**"





# Requirements Summary

# Requirements Summary

- **Ease of application integration**
- **Enterprise scalability**
- **Security**
- **Flexible policies for resource management**
- **Data Management**
- **Meta Data Management**
- **Integration with IT infrastructure, standards**

# Ease of Application Integration

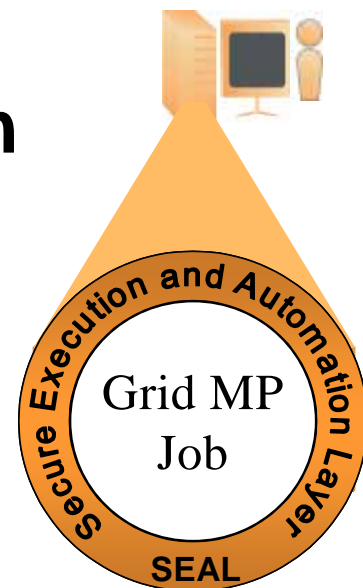
- **Needs to be quick and easy**
- **Code runs “as is”**
  - No source code modifications, re-linking, etc.
- **End-user transparency**
  - No changes for end-users interacting with the application
- **Adherence to open standards**
  - Example: web services technologies such as SOAP, XML, WSDL
- **Facilitate application license management**
- **Configuration management – version control**

# Enterprise Scalability

- **Number of Nodes**
  - Tens of thousands
- **Ease of expanding/contracting the nodes**
  - No manual intervention required
  - Editing configuration files, re-starting services, ....
- **Number of jobs**
  - A job may represent a complex workflow
- **Number of users**
- **Extending across a multi-location, multi-national enterprise**

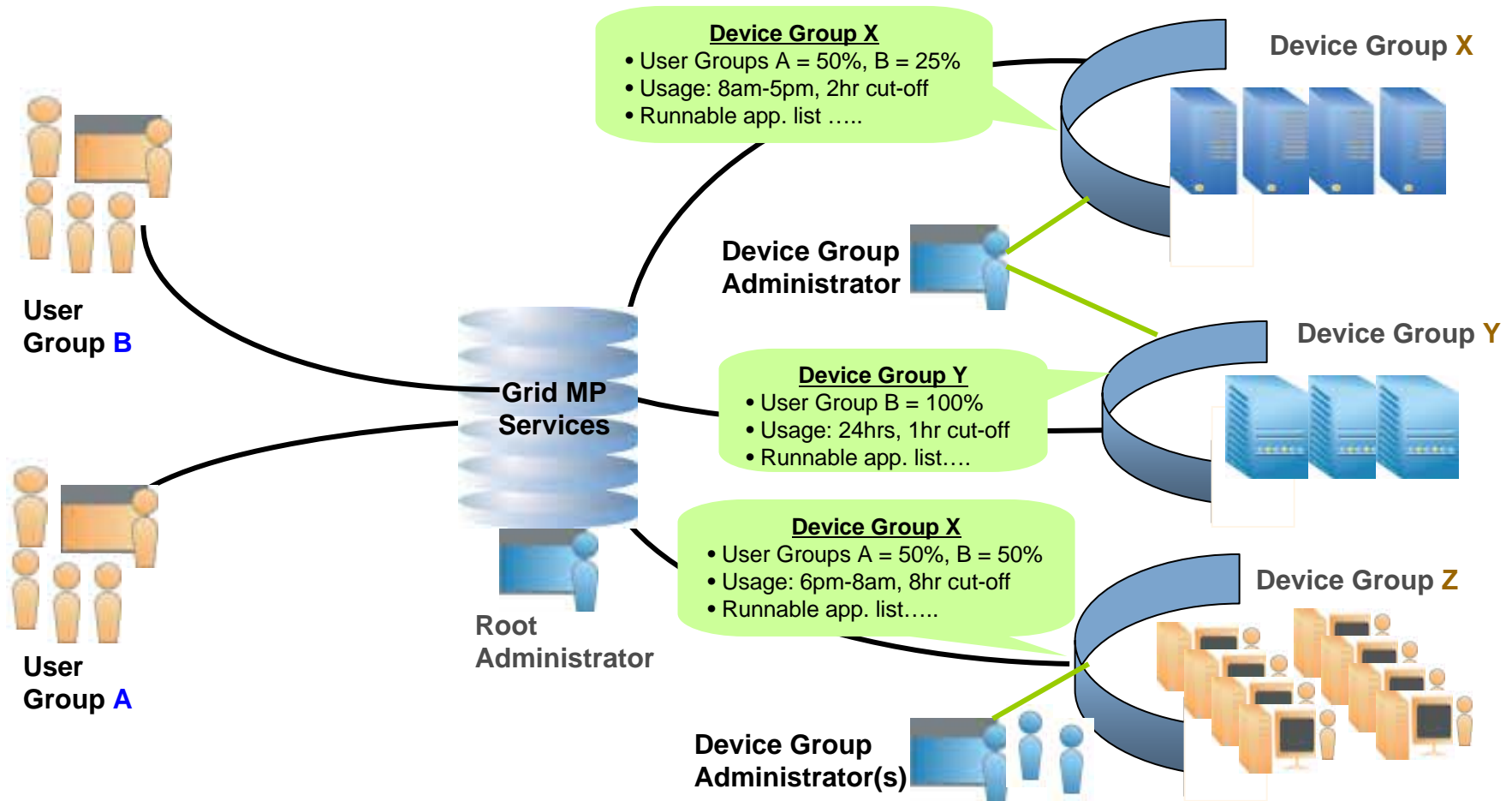
# Security

- Focus on grid user authentication only is not sufficient
- Data Protection – automatic encryption
- Data Integrity – tamper detection
- Protecting primary uses in non-dedicated environments
- Code Integrity



# Flexible policies for resource management

## *“Go ahead, hug your server!”*



# Data Management

- **Transparent data transfer**
  - Data, code needs to be migrated to the compute nodes
- **Managing application data in wide-area grid environment**
  - Many companies have low bandwidth between sites especially across the ocean
- **Optimizing data transfers**
  - Selectively constrain job placement
- **Automatic data compression**
- **Data Affinity Scheduling**
  - Caching data and scheduling based on cache contents

# Meta Data Management

- **Needs to be persistent**
- **Information for accounting, chargeback**
- **Data warehouse-based grid reporting capabilities**
  - “You cannot manage what you cannot measure”
- **Need insights for capacity planning**
- **Audit, Traceability**
- **Verifiable execution**



# Integration with IT infrastructure, standards

- **Integrate with existing standard tools for configuration management**
- **Interoperate with commercial user authentication systems**
- **Application and network management, monitoring tools**
- **Integrating with image build, qualification**
- **IT standards for 3<sup>rd</sup> party tools, documentation, processes**
- **Reliability, Availability**
  - No single points of failure in the system

# Conclusions

- **Make it easy for end users**
  - Hide complexity of “grid”
- **Reduce management overhead for IT**
- **Ease of application integration**
- **Ease of deployment both applications and management software**
- **Reliability for mission critical applications**