Next Generation Infrastructure Software for Hybrid Cloud HPC
About Nimbix

• Nimbix is the pioneer of purpose-built cloud computing for machine learning, AI and HPC applications
• Powered by JARVICE™, the Nimbix Cloud provides high-performance software as a service, dramatically speeding up data processing for Energy, Life Sciences, Manufacturing, Media and Oil & Gas software applications
• Its JARVICE™ XE Enterprise HPC platform brings the power of JARVICE™ and the Nimbix Cloud to any on-premises cluster or hybrid cloud environment
• 2019 HPCWire Readers' Choice Award winner for Best Use of High-Performance Data Analytics & Artificial Intelligence
• Recently named the Official Compute Partner of the US Sailing Team
• Trusted by Fortune 500 companies and world-class partners
HPC demand is accelerating

- From traditional supercomputers and clusters to HPC in the Cloud
- Enterprises adding new workloads, including deep learning/machine learning which demand HPC and accelerated infrastructure
- Increasing complexity to consume optimized software when it's needed
Traditional Approach of Running HPC Applications

- Traditional HPC Clusters
  - Attach to local high-performance storage
  - HPC workload managers/schedulers: Slurm, LSF, SGE, PBS…
- User scripts, jump nodes, in-house workflow automation, customized environments

![Diagram of HPC environment]

- Head node-scheduler
- Remote access
- Interconnection network
- Computing nodes
- Parallel task
- Data storage
- Applications
New Challenges to Traditional HPC

• Scaling Beyond Single Cluster, Single Location
• Replicating applications and workflows beyond single cluster
• Performance management on different infrastructures
• Increased complexity of multiple hardware architectures, multiple software stacks
• Converging with Kubernetes for AI/Deep Learning or traditional IT?
Early years of Cloud Computing for HPC

- For the first decade, cloud was completely impractical for HPC
  - Highly inadequate compute resources
  - Very complex to set up
  - Cost a fortune for attempting to run real HPC apps
  - Data security and data movement challenges created friction

- On Premises workflows & applications highly customized and not easily portable to a cloud environment
Growing heterogeneity in cloud resources
  - New CPUs: ARM, AMD, Power
  - Accelerators: GPUs, FPGAs, AI Chips

Higher performance infrastructure and interconnects: InfiniBand, bare metal

Still complex to consume and engineer in house workflows without being tethered to single provider among hyperscalers
Desired State: **Hybrid On-Prem/Cloud for Enterprise HPC**

- Balance and Optimize economics of on prem clusters and cloud use
- Preserve existing applications and workflows
- Limit customizations for any given infrastructure or cloud platform – make it adaptable
- Optimize for policy-driven computing
- Full enterprise reporting across all resources
Enterprise Apps increasingly need to be adaptable and available on any infrastructure and multiple cloud providers.
Proliferation of Containers

• “Inter-modal” method for packaging, distributing, and deploying applications and dependencies onto arbitrary infrastructure
  • Consistency and integrity across platforms
  • Repeatable mechanisms
• Displacing traditional VMs for application image management
• Every major platform provider supports Docker containers
• Enables applications to take advantage of bare-metal infrastructure capabilities for high performance workloads
Containerized HPC Application Comparison

**Docker**
- Container-native capable
- Multi-tenant
- Unlimited, diverse use cases

- **Head** node
  - Container 0
  - Application code
  - mpi.run ...
  - MPI libraries

- **Node 1**
  - Container 1
  - Application code
  - MPI libraries

- **Node n**
  - Container n
  - Application code
  - MPI libraries

**Secure Container network/fabric**

**Singularity**
- Traditional/monolithic
- Single-tenant
- Limited or related use cases

- **Head** node
  - Container 0
  - Application code
  - MPI libraries

- **Node 1**
  - Container 1
  - Application code
  - MPI libraries

- **Node n**
  - Container n
  - Application code

**Shared Host network/fabric**

*Low-latency fabric (e.g. InfiniBand, RoCE), requires HPC-capable container platform, such as JARVICE*
• Brings industry leading HPC container technology to your datacenter, your cloud or any cloud, anywhere in the world

• Delivers and synchronizes entire HPC/ML application ecosystems to your clusters via HyperHub™

• Rapidly deploys and synchronizes your applications and workflows to your HPC environments around the world

• Enables rapid adoption of new technologies and new infrastructure even before hyperscale clouds make them available
HPC on Kubernetes with JARVICE XE
JARVICE XE Features

• Unified Kubernetes infrastructure
• HyperHub application ecosystem syncs commercial, custom or in-house applications and workflows globally
• Enterprise security integrates with Active Directory and SAML for seamless, secure authentication and authorization
• Comprehensive enterprise analytics and reporting across all HPC workflows and jobs across all global clusters
• Seamless global updates of both platform and applications
• Secure containerized architecture
• Compatible with every major cloud provider and any HPC compute infrastructure
Simplified User Experience
• Simple point-click-run workflows on any infrastructure

Simplified Administration for HPC Admins and Enterprise IT
• Unified SaaS for HPC and Deep Learning
• Unified platform for multi-cloud, multi-datacenter deployments

Streamlined Application Distribution and Deployment
• Platform-as-a-Service (PaaS) continuous integration and deployment for compute intensive workflows using your in-house algorithms, applications, or customization of commercial applications
• Automatic synchronization with HyperHub™

Reduced Infrastructure Complexity
• Unified infrastructure layer with Kubernetes
Next Gen Platform Software Optimizing HPC Costs

- Replace alternative common solution stacks that combine commercial schedulers
- Deploy on open source Kubernetes software that works with existing Linux infrastructure, eliminating the need for expensive HPC-specific cluster management or infrastructure management solutions
- Reduce/Eliminate the need for dedicated HPC application engineering resources, since application management is self-service and does not require "installing" anything on servers, everything is containerized
- Allow IT/HPC admins to focus resources on more general projects rather than having to specialize on HPC management
- Improve user experience and access from anywhere means fewer helpdesk calls and less management of end user clients, a web browser is all that is needed
- Seamlessly burst to other infrastructure when needed, avoiding the need to procure more HPC hardware to add temporary, project-based, or seasonal capacity to core clusters
Key Attribute Considerations for **Next Gen HPC** Software Platforms

<table>
<thead>
<tr>
<th>Desired State</th>
<th>Open Marketplace</th>
<th>Platform (PaaS)</th>
<th>Hybrid/Multi-Cloud</th>
<th>Accelerator-Ready</th>
<th>Ease-of-Use / Time to Deploy</th>
<th>Ecosystem</th>
<th>Performance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Traditional Scheduler Solutions</td>
<td><img src="Image" alt="Circle" /></td>
<td><img src="Image" alt="Circle" /></td>
<td><img src="Image" alt="Circle" /></td>
<td><img src="Image" alt="Circle" /></td>
<td><img src="Image" alt="Circle" /></td>
<td><img src="Image" alt="Circle" /></td>
<td></td>
</tr>
<tr>
<td>DIY with Cloud</td>
<td><img src="Image" alt="Circle" /></td>
<td><img src="Image" alt="Circle" /></td>
<td><img src="Image" alt="Circle" /></td>
<td><img src="Image" alt="Circle" /></td>
<td><img src="Image" alt="Circle" /></td>
<td><img src="Image" alt="Circle" /></td>
<td></td>
</tr>
<tr>
<td>DIY In-House HPC</td>
<td><img src="Image" alt="Circle" /></td>
<td><img src="Image" alt="Circle" /></td>
<td><img src="Image" alt="Circle" /></td>
<td><img src="Image" alt="Circle" /></td>
<td><img src="Image" alt="Circle" /></td>
<td><img src="Image" alt="Circle" /></td>
<td></td>
</tr>
</tbody>
</table>
HyperHub™ Application Marketplace

- Distribute, deploy, and run applications and workflows globally on any cloud or on premises datacenter

- Built-in support for autoscaling apps for high performance clusters

- Native accelerator support for both GPU and FPGA infrastructure

- HPC and supercomputing applications as point-and-click workflows

- Automatic synchronization of marketplace to any JARVICE XE deployment
HyperHub™ Ecosystem

- Offers a curated catalog of Simulation, AI/ML/DL software optimized for HPC
- Combined with JARVICE XE enables users to access the entire catalog of applications with virtually unlimited supercomputing power on any infrastructure
- Customers can customize the compute catalog with their own workflows and applications
Recommended Reading

- Download the White Paper: HPC on Kubernetes
  https://www.nimbix.net/case-studies-white-papers
Contact Us

THANK YOU!

- Experience the benefits of JARVICE XE today with a complementary trial
- Reach out to your Nimbix or Partner representative for information on Nimbix Cloud and JARVICE XE availability: sales@nimbix.net
- Visit us at www.nimbix.net

linkedin.com/company/nimbix  facebook.com/nimbix  twitter.com/nimbix