

**WASHINGTON STATE UNIVERSITY**

**Institutional Research Computing at WSU:  
Implementing a community-based approach**

**Exploratory Workshop on the  
Role of High-Performance Computing in the Pacific Northwest**

**University of Washington**

**July 27, 2015**

**Christian Mailhot**

**Professor, College of Arts and Sciences**

**[christian.mailhot@wsu.edu](mailto:christian.mailhot@wsu.edu)**



**WASHINGTON STATE  
UNIVERSITY**  
*World Class. Face to Face.*

# Outline – Institutional Research Computing at WSU

- **Summary of science drivers:**

- Application domains
- Scientific computing needs

- **Strategy for institutional research computing:**

- Cyber-infrastructure
- HPC platform – research computing cluster
- Research computing user support – software implementation, development, and optimization
- Software inventory across application domains

- **Condominium computing as a community-based approach: Management model**

- Governance model
- Investment model – acquisition strategy

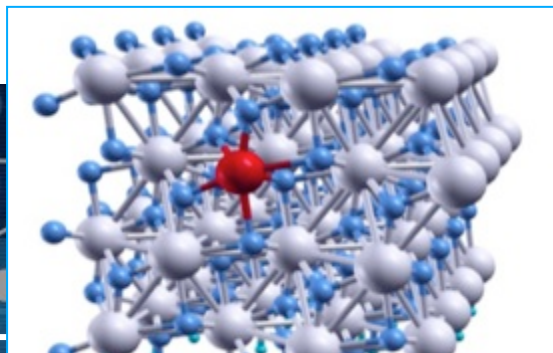
- **Alignment with WSU long-term research computing strategy:**

- Faculty cluster hire in scientific and data computing
- Build a unified research computing initiative by integrating needs and resources across WSU campuses

# Representative WSU science drivers and application domains



Genomics, Genetics, Bioinformatics, Agriculture, Health Sciences



Physics, Materials Science and Engineering, Chemistry and Biochemistry

**WSU Institutional Research Computing:**  
Building the capacity to lead with a focus on applications



Smart Energy Grid



Environmental and Atmospheric Research

## Application domains at WSU

**Genomics, Genetics, Bioinformatics, Agriculture:** Evolutionary genomics, biomedical genomics, crop genomics, breeding research, software platform for next-generation data analysis and sharing.

**Physics, Materials Science and Engineering, Chemistry and Biochemistry:** Materials Genomics, computational design of materials, materials for clean energy, materials in extreme environments, actinide chemistry, catalysis, nuclear theory, computational astrophysics.

**Atmospheric and Environmental Research:** Air quality forecasting, numerical weather prediction, regional-scale earth system modeling, watershed integrated systems dynamics modeling.

**Smart Energy Grid:** Power system analysis, control enhancement, demand management, cyber-physical security to power infrastructure.

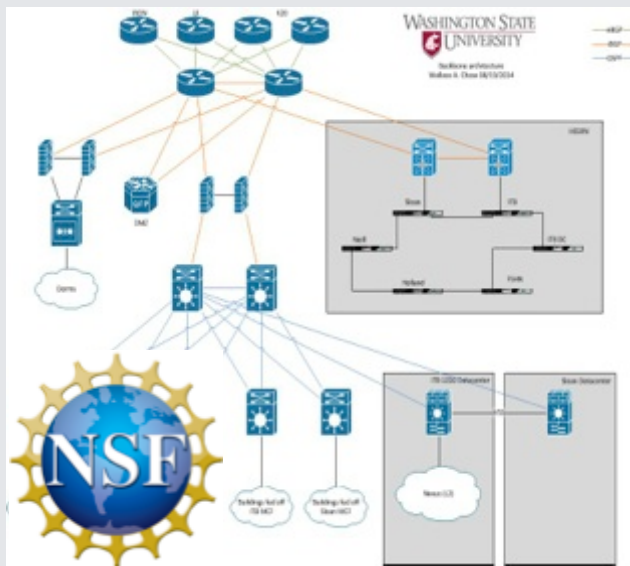
**Health Sciences:** Biomedical genomics, systems pharmacology.

**Education and training:** Computational science, computer science, data science, AI, bioinformatics, HPC training.



# The WSU institutional research computing strategic plan: Cyber-infrastructure, computer platforms, software user support

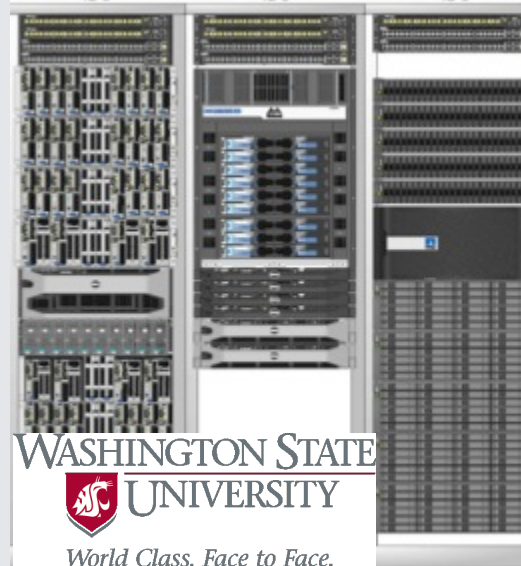
## Improving the WSU cyber-infrastructure: High-speed research network – High-performance data storage / transfer



- **2014 NSF CC\*IIE award (\$500 K):** High Speed Scalable Research Core (HSSRC) – Science DMZ
- **2015 NSF MRI award (\$500 K – including cost share)(\*):** High-performance data storage/transfer between WSU Pullman and Spokane campuses.

(\* ) Recommended for funding

## Condominium computing: Integrated and scalable approach



- Implement condominium computing for institutional research computing
- Start with a “**pilot**” **compute cluster** that can be expanded by the research community
- Implement a business model to re-capitalize the infrastructure

## Establishing a user support group for software implementation



- Establish an applications-focused **research computing user support group** to accelerate the installation, development, and optimization of application software tools
- Research associate for user support group being hired
- Software inventory in progress

## Institutional research computing resources: Platforms managed by central IT

### Current HPC platform IBM I-Dataplex (2011)



#### — Compute nodes:

- 164 CPU nodes (12 cores – 24 GB)

#### — Large memory nodes:

- 3 CPU nodes (32 cores – 512 GB)

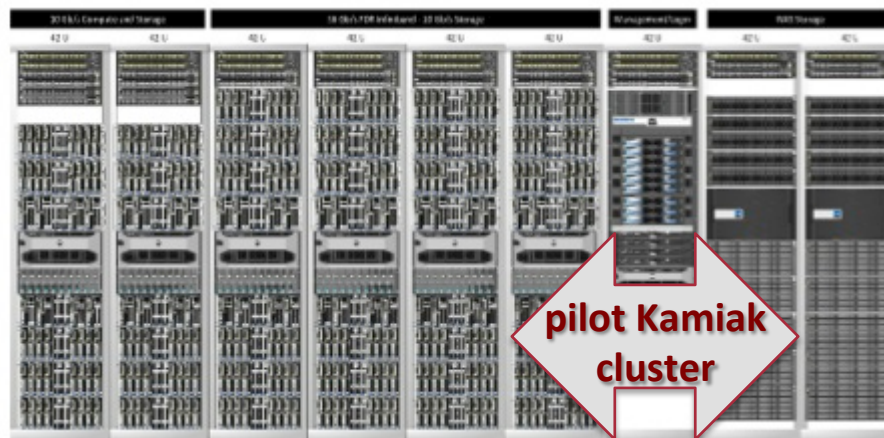
#### — Storage:

- No physical local disk at compute nodes

#### — Network:

- Infiniband switch (1)
- 40 Gb switch (1)

### WSU Kamiak (pilot) Cluster (2015 +)



#### — Compute nodes:

- 32 CPU nodes (20 cores – 256 GB / 512 GB)
- 2 NVIDIA GPU node
- 1 Phi (Intel Xeon) GPU node

#### — Large memory nodes:

- 2 TB RAM Server (60 cores)

#### — Storage:

- NetApp File Storage (633 TB)

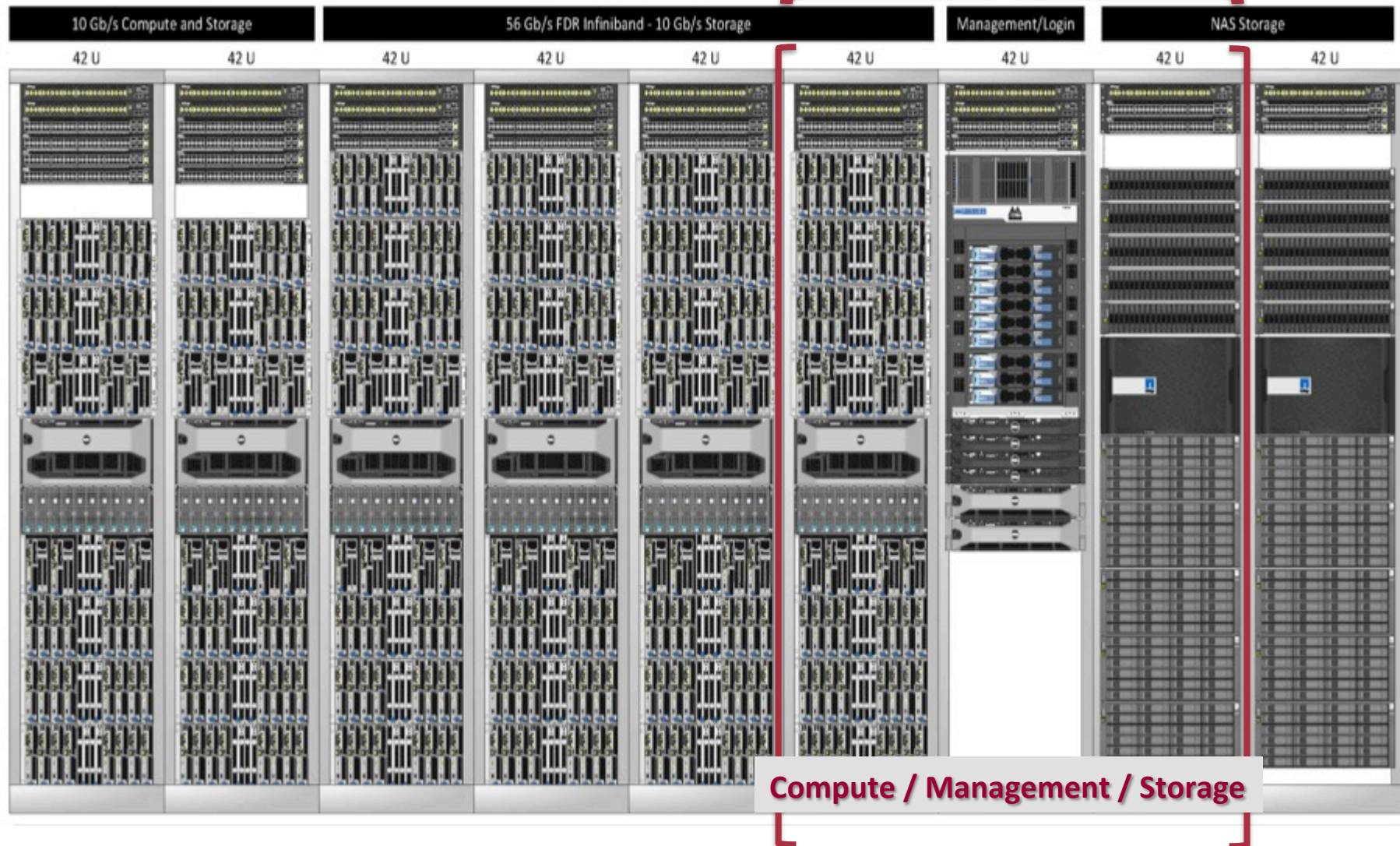
#### — Network:

- Infiniband switch (1)
- 40 Gb switch (1)
- 10 Gb switches for network storage (4)
- 10 GbE switches for network storage (3)



**The WSU full-size condominium Kamiak cluster (phase 1):  
9-rack system: Equipment and research grants; start-up funds; and other contributions  
from faculty, researchers and academic units**

**WSU pilot Kamiak cluster**



## Inventory of application software at WSU: Exploring site licenses for widely used software

Application domain	Software
<ul style="list-style-type: none"> <li>— Genomics</li> <li>— Proteomics</li> <li>— Bioinformatics</li> </ul>	<ul style="list-style-type: none"> <li>— Over 150 software applications</li> <li>— Most applications are open source</li> </ul>
<ul style="list-style-type: none"> <li>— Materials science</li> <li>— Mechanical engineering</li> <li>— Electronic structure</li> <li>— Chemistry</li> </ul>	<ul style="list-style-type: none"> <li>— VASP, Quantum Espresso, Abinit, Siesta, Wien2k</li> <li>— LAMMPS</li> <li>— CASTEP, Materials Studio</li> <li>— NWChem, Gaussian, Molpro, Dirac, CFour, DL_POLY</li> </ul>
<ul style="list-style-type: none"> <li>— Atmospheric research</li> <li>— Civil and Environmental Engineering</li> </ul>	<ul style="list-style-type: none"> <li>— Portland Group compilers</li> <li>— NCL (NCAR Command Language)</li> <li>— ABAQUS, ANSYS (Finite elements)</li> <li>— Fluid dynamics</li> <li>— Finite elements analysis</li> </ul>
<ul style="list-style-type: none"> <li>— Math and statistics</li> <li>— Optimization</li> <li>— Simulation</li> <li>— Model-based design</li> </ul>	<ul style="list-style-type: none"> <li>— MATLAB (MathWorks)</li> </ul>

## WSU is implementing a condominium approach to institutional research computing

What ?	Who ?	How ?	Comments
<b>Sponsors</b>	<ul style="list-style-type: none"> <li>— Colleges</li> <li>— Academic units</li> <li>— Office of Research</li> </ul>	<b>IT and research computing staffing:</b> <ul style="list-style-type: none"> <li>— Systems administrator</li> <li>— User support for research computing</li> </ul>	Possible contributions to cyber-infrastructure
<b>Investors</b>	Faculty and researchers who require predictable computational availability	<ul style="list-style-type: none"> <li>— Purchase own equipment (compute nodes, storage, etc.)</li> <li>— WSU/ITS purchases the nodes and deploy them in the shared infrastructure and operate them for a fixed number of years</li> </ul>	<ul style="list-style-type: none"> <li>— Once installed, purchased nodes become part of the Kamiak cluster</li> <li>— Cost for a node is price of equipment + markup for IT systems administration and user support</li> </ul>
<b>General users</b>	Entire WSU community	Sponsored by their administrative College	<ul style="list-style-type: none"> <li>— Unused compute cycles in the condominium are available for general users</li> <li>— Access to “backfill queue” can be preempted by investors’ priority access</li> </ul>
<b>Institution</b>	<ul style="list-style-type: none"> <li>— Office of Provost</li> <li>— Office of Research</li> <li>— Office of Finance</li> <li>— WSU/ITS</li> </ul>	<b>Physical infrastructure:</b> <ul style="list-style-type: none"> <li>— Equipment room space</li> <li>— Power, cooling, etc.</li> <li>— Racks</li> </ul>	Possible contributions to cyber-infrastructure



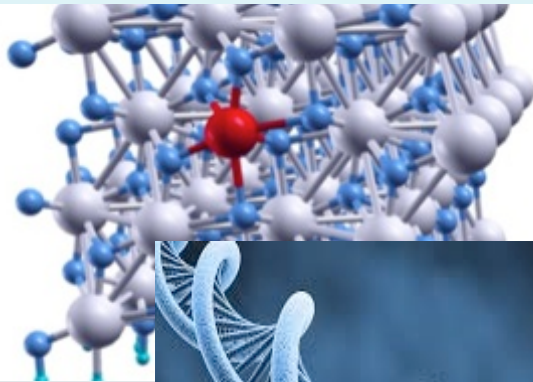
## **WSU is committing resources to establish a user support group for application software implementation, optimization, and development**

- **Establishment of a software user support group: “IT Research Computing Consultant”**
  - Focus on research computing
  - Provide assistance in software installation, development, and optimization
  - Broad spectrum of application domains:
    - Materials science and engineering
    - Chemistry and chemical engineering
    - Bioinformatics
    - Genomics
    - Atmospheric research
  - Parallel scientific computing
  - Installation and management of software libraries
  - Development of documentation and training material for the effective use of institutional HPC resources
- **Institutional support:**
  - Institutional support from Colleges and the Office of the VPR (1 FTE)

# Implementing an institutional strategy to advance research computing “at-scale” across the WSU system

## Cluster faculty hires in scientific and data computing (2015)

### Computational materials science



### Bioinformatics

#### • 2015 – 2016 cluster faculty hires (4):

- Computational materials science and engineering (Beckman)
- Computational materials physics (McMahon)
- Bioinformatics (Ficklin)
- Computational astrophysics (TBD)

## Connecting the WSU research computing community



- Implementing community computing across the university system to support the WSU research computing strategic plan.
- Strategically enhancing the cyber-infrastructure by extramural funds and institutional investments.