

# The Zero-Force MPI Toolkit— Toward Tractable Toolkits for HPC

Presented by

**Magdalena Slawinska**  
**Dawid Kurzyniec**  
**Jaroslaw Slawinski**  
**Vaidy Sunderam**

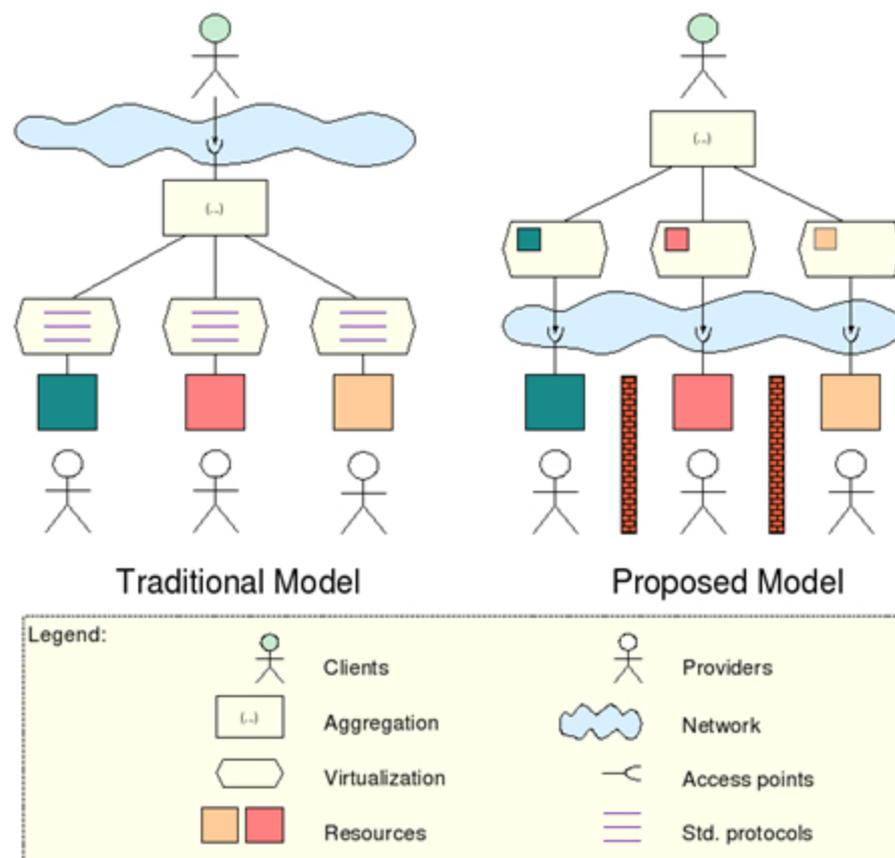
Emory University



# The Zero-Force MPI Toolkit—Toward tractable toolkits for high-performance computing

## Goals

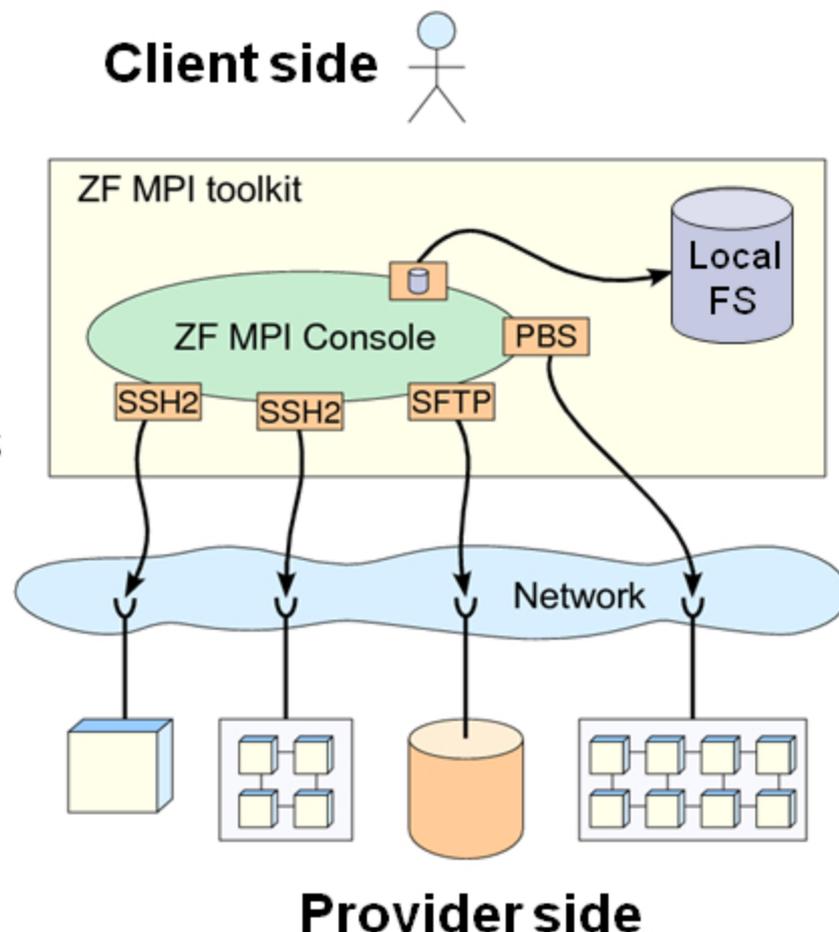
- **New approach to resource sharing**
  - Aggregation and virtualization resources at the client side
- **Automating tasks**
  - MPI environment configuration
  - Uploading and compiling computational applications
  - Staging input data/result collection



# The Zero-Force MPI Toolkit— Toward tractable toolkits for HPC

## Architecture

- **Client side: Interactive console**
  - Unified and coherent interface
  - Resource virtualization
  - Mediators—“service-drivers”
- **Provider side: Access daemons**
  - E.g., sshd, ftpd
- **Prototype implementation**
  - FT-MPI
  - Java
  - JSch (SSH2 implementation)



# The Zero-Force MPI Toolkit— Toward tractable toolkits for HPC

## ZF-MPI console

**DVM  
assembly**

**Add computational resource to DVM**

- Upload FT-MPI
- Decompress sources
- Install
- Launch FT-MPI daemons

**Data sync**

**Synchronize data, source files across  
DVM nodes**

**Compile  
and build**

**Invoke standard shell commands**

**Application  
launch**

**Run MPI programs**

# The Zero-Force MPI Toolkit— Toward tractable toolkits for HPC

## *Stop FT-MPI VM*

```
zf-mpi> add ft_mpi joe@{lab6a,lab6b,lab6c,lab6d,compute}
zf-mpi> ft_mpi setNS compute
zf-mpi> ft_mpi add lab6a,lab6b,lab6c,lab6d,compute
zf-mpi> sync ~/NPB3.2.1/NPB3.2-MPI ~/zf-mpi/
zf-mpi> cd ~/zf-mpi/NPB3.2-MPI
zf-mpi> make bt NPROCS=4 CLASS=B
zf-mpi> mv bin/bt.B.4 $HARNESS_BIN_DIR/$HARNESS_ARCH/
zf-mpi> ft_mpi ftmpirun compute -np 4 -o bt.B.4 > log
zf-mpi> cat log | grep "Time in seconds"
zf-mpi> ft_mpi console haltall
```



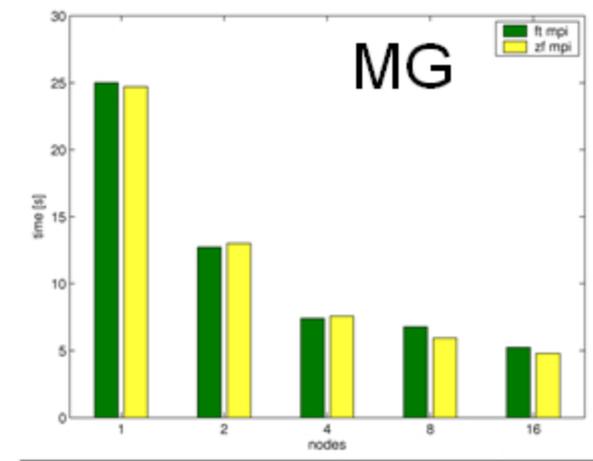
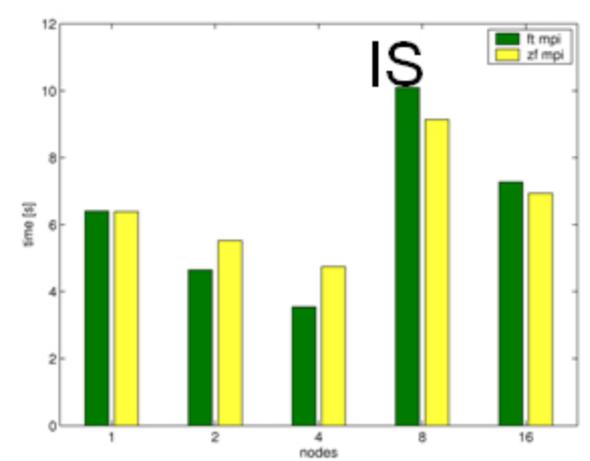
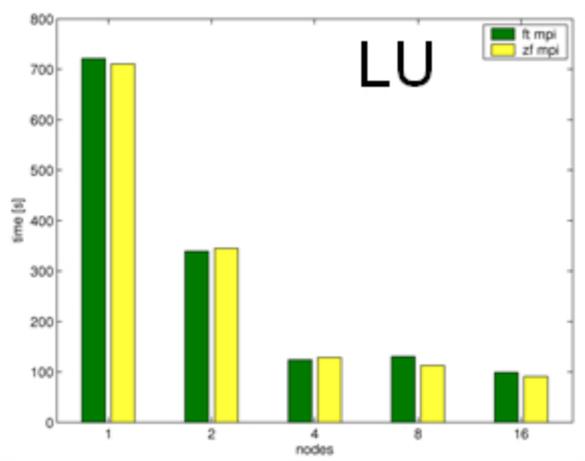
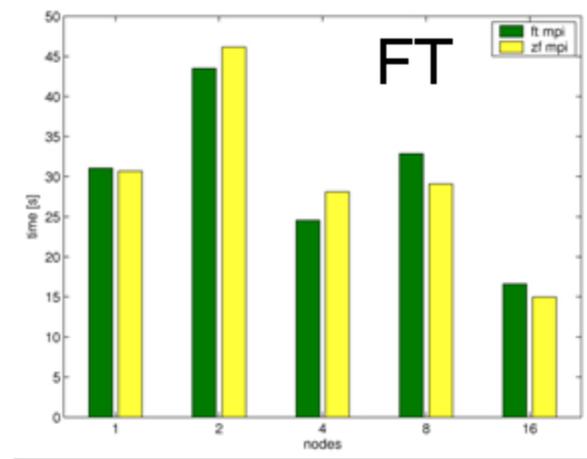
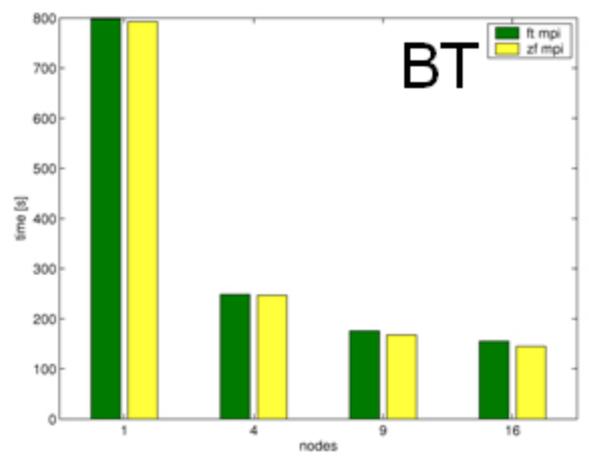
# The Zero-Force MPI Toolkit— Toward tractable toolkits for HPC

## Experimental evaluation

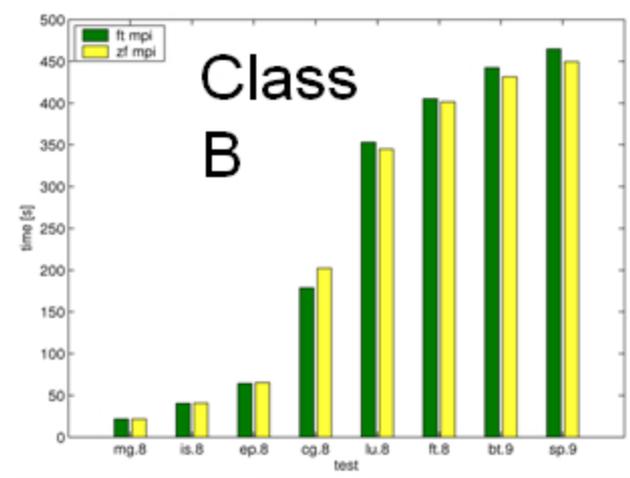
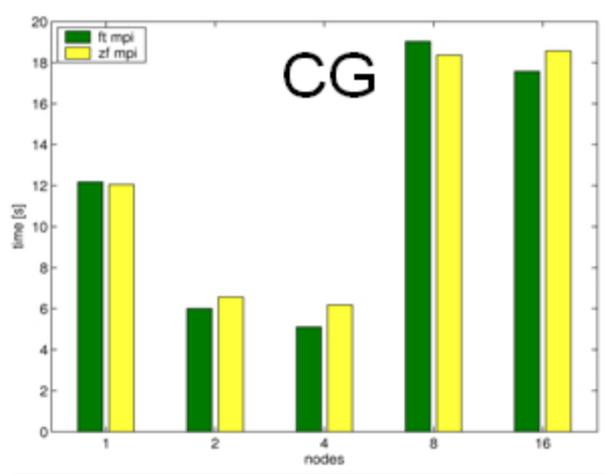
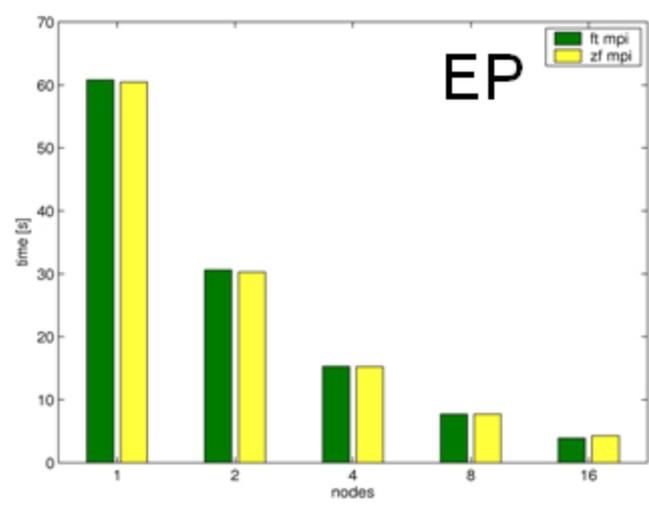
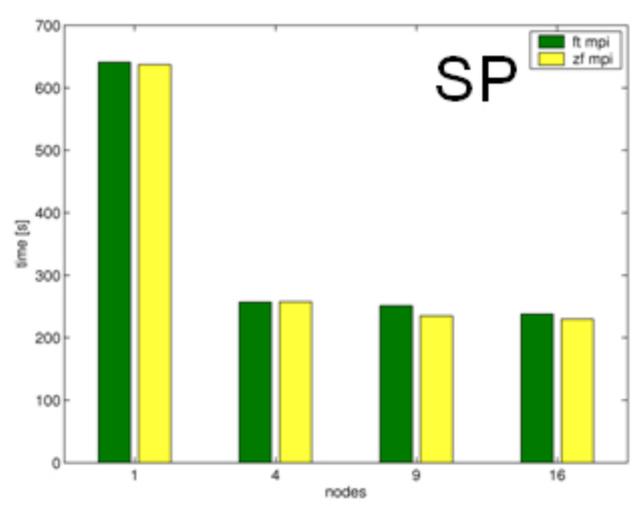
- **NAS Parallel Benchmarks 3.2 for MPI**
- **Linux/i86**
  - **Pentium 4 (2.4–2.8 GHz), 1 GB RAM**
  - **Linux Mandriva 2006 (kernel 2.6.12-12)**
- **Solaris (non-GNU-based UNIX)**
  - **Sun Blade 2500, UltraSPARC-III, 1280 MHz**
  - **1 MB cache memory per processor, 2 GB RAM**
  - **Connected directly to 100 Mbit HP network switches**
  - **SunOS 5.10**
- **Homogeneous cluster: Class A, up to 16 processes**
- **Heterogeneous cluster: Class B, 8, 9 processes**



# The Zero-Force MPI Toolkit— Toward tractable toolkits for HPC



# The Zero-Force MPI Toolkit— Toward tractable toolkits for HPC

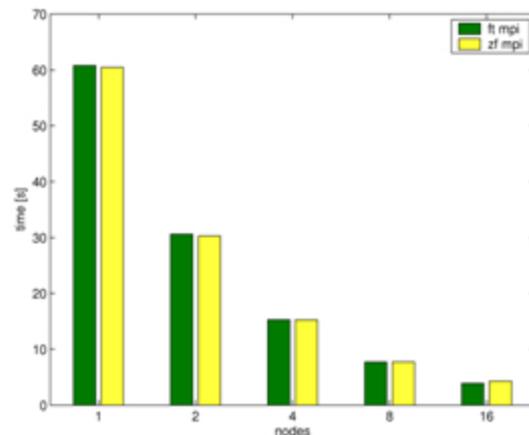
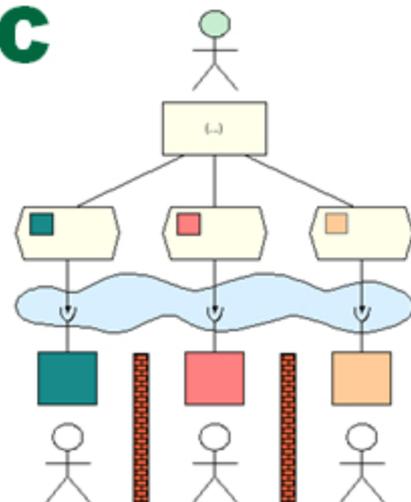
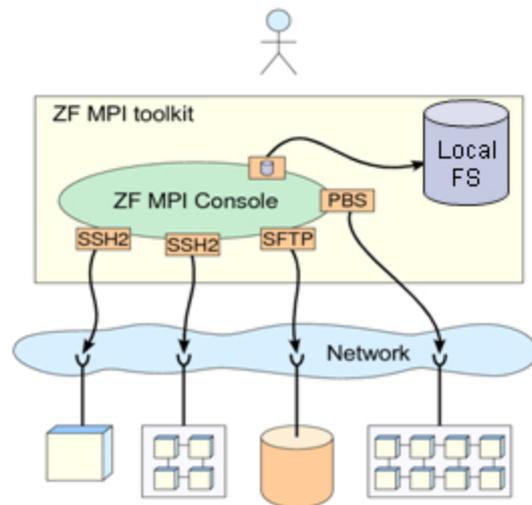


# The Zero-Force MPI Toolkit— Toward tractable toolkits for HPC

## Summary

### ZF-MPI

- Implements the new resource sharing model
- Demonstrates feasibility of the run-not-install approach
- Significantly reduces efforts such as deployment and execution
- Does not affect performance of MPI applications



# Contacts

## **Magdalena Slawinska**

Emory University  
magg@mathcs.emory.edu

## **Jaroslav Slawinski**

Emory University  
jaross@mathcs.emory.edu

## **Vaidy Sunderam**

Emory University  
vss@mathcs.emory.edu

