



UNCLASSIFIED: Distribution Statement A.
Approved for public release



TECHNOLOGY DRIVEN. WARFIGHTER FOCUSED.

Computational Fluids Domain Reduction to a Simplified Fluid Network

19 April 2012

Project Lead:
Rob E Smith
rob.e.smith@us.army.mil
586-282-4121

Report Documentation Page

Form Approved
OMB No. 0704-0188

Public reporting burden for the collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to a penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number.

1. REPORT DATE 09 JUL 2012		2. REPORT TYPE Briefing Charts		3. DATES COVERED 09-07-2012 to 09-07-2012	
4. TITLE AND SUBTITLE Computational Fluids Domain Reduction to a Simplified Fluid Network				5a. CONTRACT NUMBER	
				5b. GRANT NUMBER	
				5c. PROGRAM ELEMENT NUMBER	
6. AUTHOR(S) Robert Smith				5d. PROJECT NUMBER	
				5e. TASK NUMBER	
				5f. WORK UNIT NUMBER	
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) U.S. Army TARDEC ,6501 E.11 Mile Rd,Warren,MI,48397-5000				8. PERFORMING ORGANIZATION REPORT NUMBER #22880	
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) U.S. Army TARDEC, 6501 E.11 Mile Rd, Warren, MI, 48397-5000				10. SPONSOR/MONITOR'S ACRONYM(S) TARDEC	
				11. SPONSOR/MONITOR'S REPORT NUMBER(S) #22880	
12. DISTRIBUTION/AVAILABILITY STATEMENT Approved for public release; distribution unlimited					
13. SUPPLEMENTARY NOTES					
14. ABSTRACT briefing charts					
15. SUBJECT TERMS					
16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT Public Release	18. NUMBER OF PAGES 81	19a. NAME OF RESPONSIBLE PERSON
a. REPORT unclassified	b. ABSTRACT unclassified	c. THIS PAGE unclassified			

Agenda



- About Rob E Smith
- Research Objectives
- Quick Introduction to Data Mining
- cfdMine Tool
- cfdMine Mahalanobis
- Validation Results
- Conclusions/ Next Steps

About Me



- Career:
 - Present: Team Lead for the CFD/Signature Modeling, TARDEC
 - 2003-2008 Survivability Specialist, General Dynamics Land Systems
 - 1999-2003 Research Engineer, ThermoAnalytics, Inc.
 - 1994-1996 Test Engineer/ Lab Tech, Whirlpool Corporation
- Education:
 - 2002 MSME Michigan Tech
 - Project: Infrared Signature Modeling of the AH-64 Apache
 - 1999 BSME Michigan Tech
- Classes Taken Specific to This Thesis:
 - Data Mining
 - Neural Networks
 - Intelligent Systems

Agenda

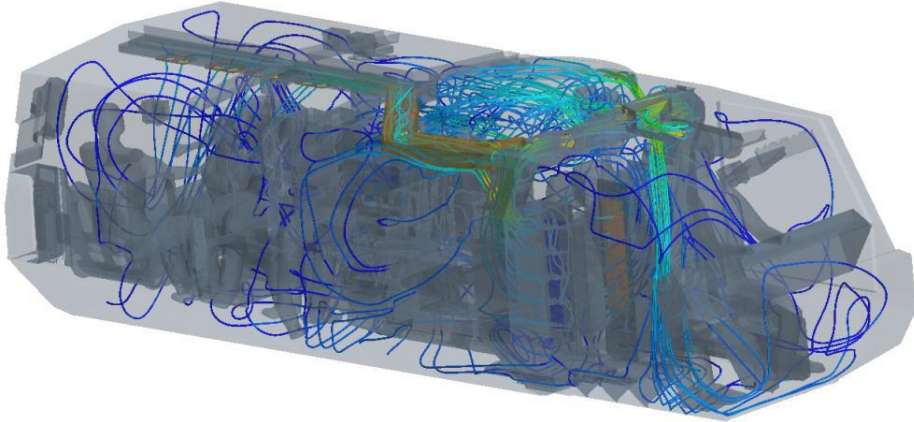


- About Rob E Smith
- Research Objectives
- Quick Introduction to Data Mining
- cfdMine Tool
- cfdMine Mahalanobis
- Validation Results
- Conclusions/ Next Steps

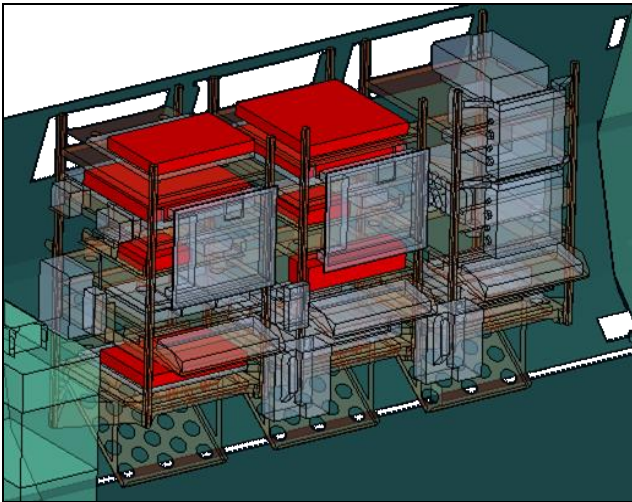
Basic Problem



Typical Military Vehicle HVAC Simulation



- CFD models of vehicle interior are ~ 15 million cells
- Solve steady state HVAC with 64 cores overnight
- Need to perform heat soaked vehicle simulation and HVAC startup
 - Would take at least a week
- Particularly interested in equipment
 - Want operational temps in less than 15 minutes

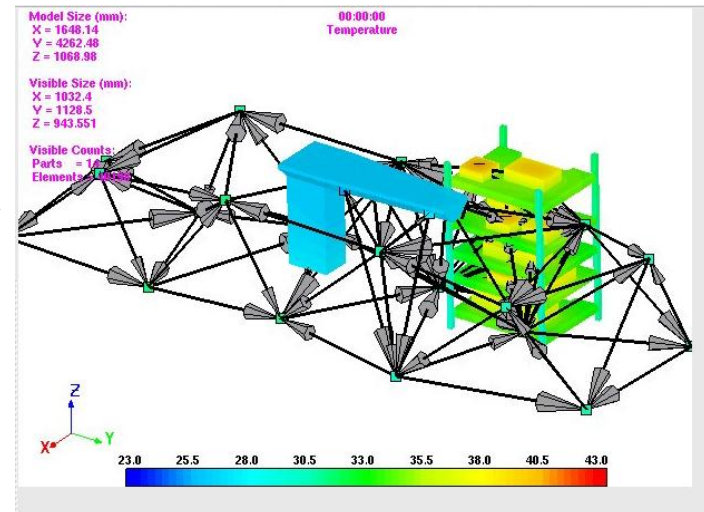
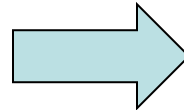
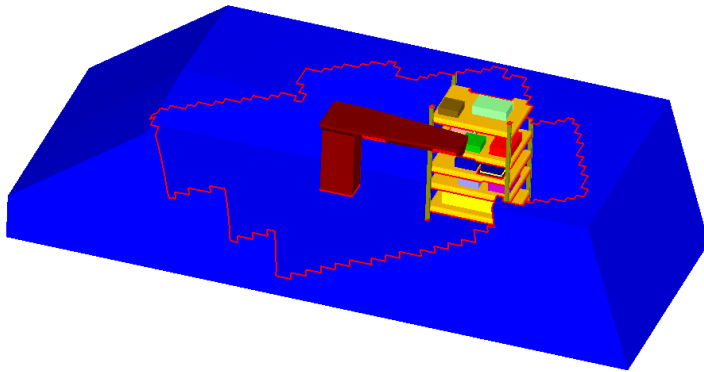


Typical Military Vehicle Equipment Rack

Research Objectives



1. Demonstrate process of clustering a steady-state CFD domain into a transient lumped fluid network simulation

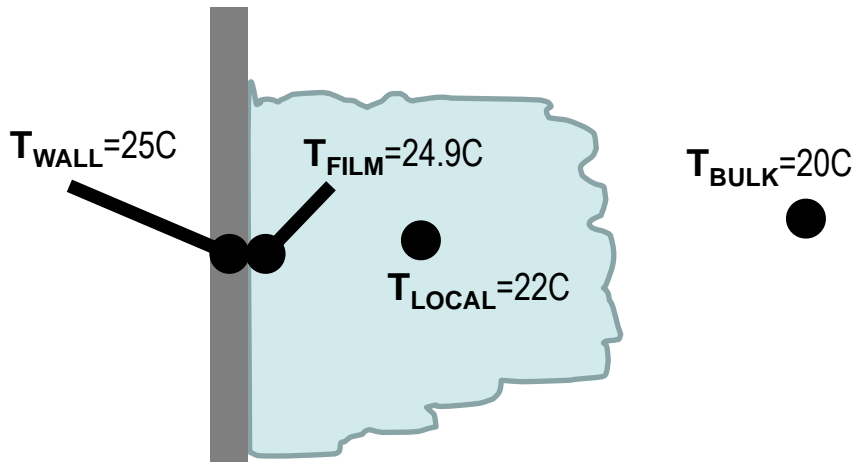


Research Objectives



1. Demonstrate process of clustering a steady-state CFD domain into a transient lumped fluid network simulation
2. Show value of remapping CFD fluid film convection to localized fluid nodes

Which H?



Remapping Equation

$$h_{film}(T_{film} - T_{wall}) = q' = h_{local}(T_{local} - T_{wall})$$

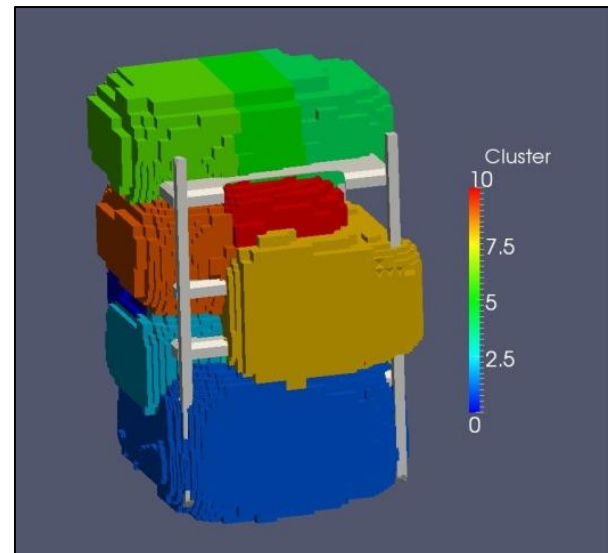
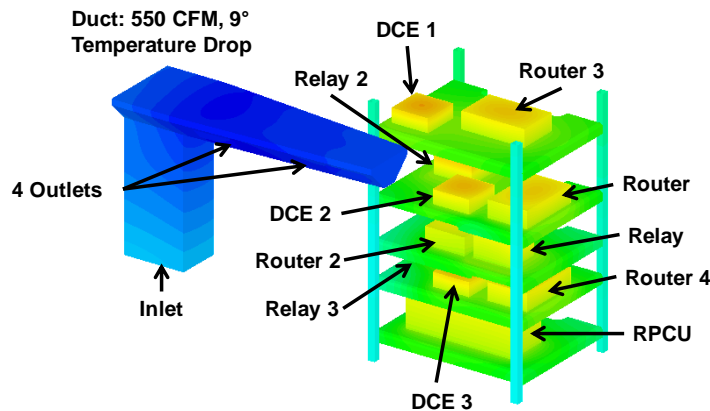
so

$$h_{local} = \frac{h_{film}(T_{film} - T_{wall})}{T_{local} - T_{wall}}$$

Research Objectives



1. Demonstrate process of clustering a steady-state CFD domain into a transient lumped fluid network simulation
2. Show value of remapping CFD fluid film convection to localized fluid nodes
3. Investigate the use of clustering to track temperatures around specific equipment or locations

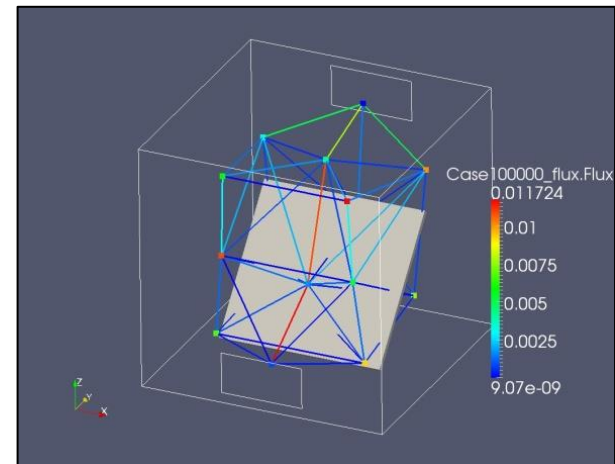
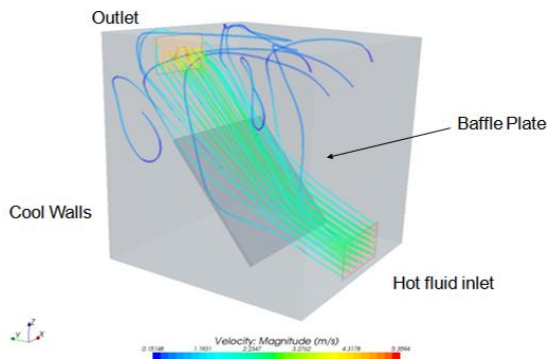


TECHNOLOGY DRIVEN. WARFIGHTER FOCUSED.

Research Objectives



1. Demonstrate process of clustering a steady-state CFD domain into a transient lumped fluid network simulation
2. Show value of remapping CFD fluid film convection to localized fluid nodes
3. Investigate the use of clustering to track temperatures around specific equipment or locations
4. Show new ways of flowfield visualization using clustering



TECHNOLOGY DRIVEN. WARFIGHTER FOCUSED.

Research Objectives



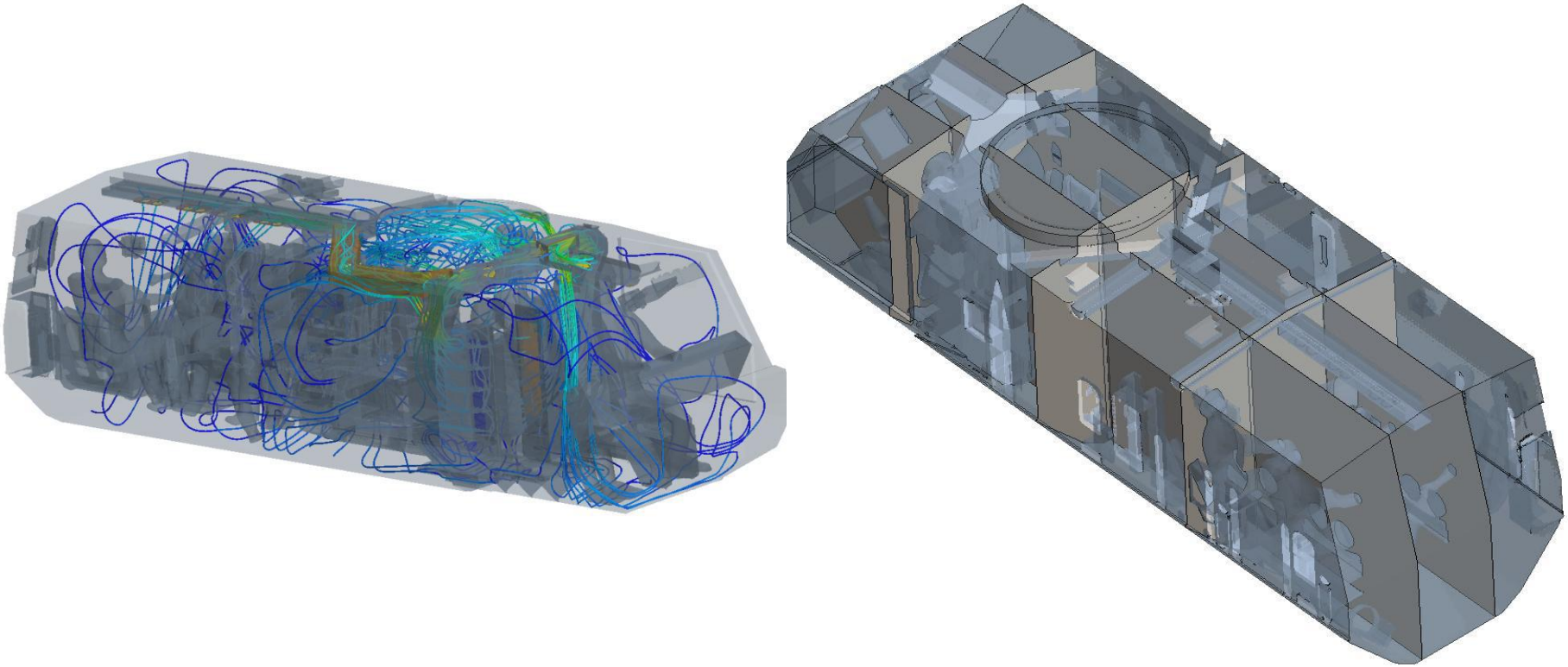
1. Demonstrate process of clustering a steady-state CFD domain into a transient lumped fluid network simulation
2. Show value of remapping CFD fluid film convection to localized fluid nodes
3. Investigate the use of clustering to track temperatures around specific equipment or locations
4. Show new ways of flowfield visualization using clustering
5. Validate subdivoluming and new clustering techniques against fully transient CFD data



TECHNOLOGY DRIVEN. WARFIGHTER FOCUSED.

UNCLASSIFIED

State of the Art: Subvoluming Approach



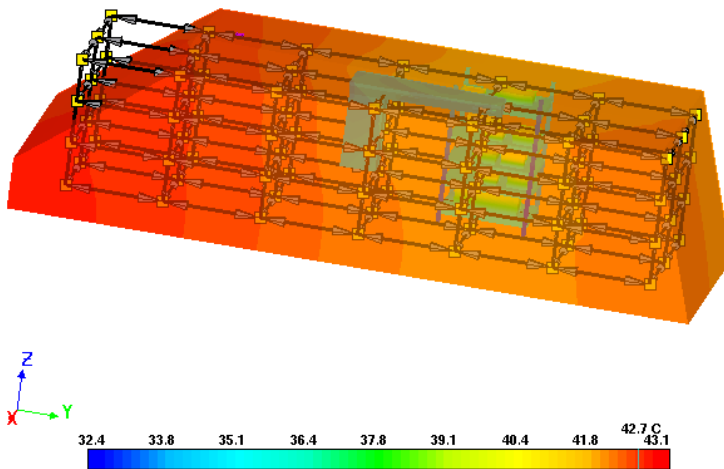
- Simple idea: cut planes through the domain and track volume fluxes across surfaces to make control volume network.
- Apply a fixed 15 w/m^2 convection coefficient that convects to closest node

TECHNOLOGY DRIVEN. WARFIGHTER FOCUSED.

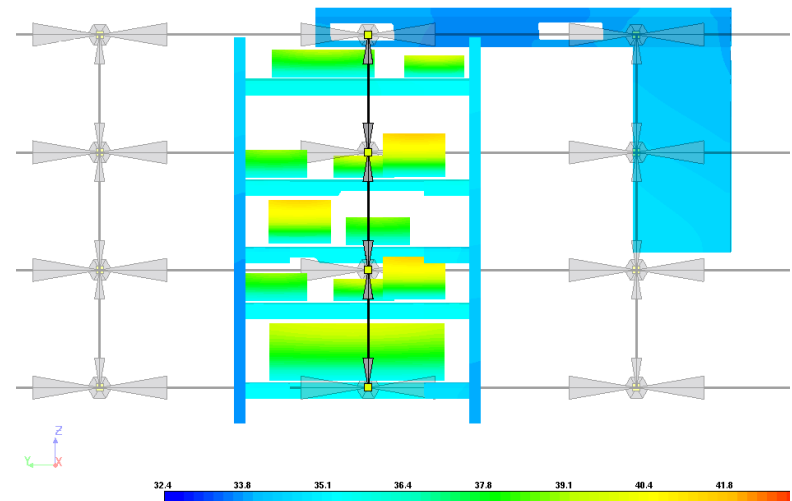
Subvolumed Model / Problems



Subvolume nodal arrangement
72 Nodes Evenly Spaced



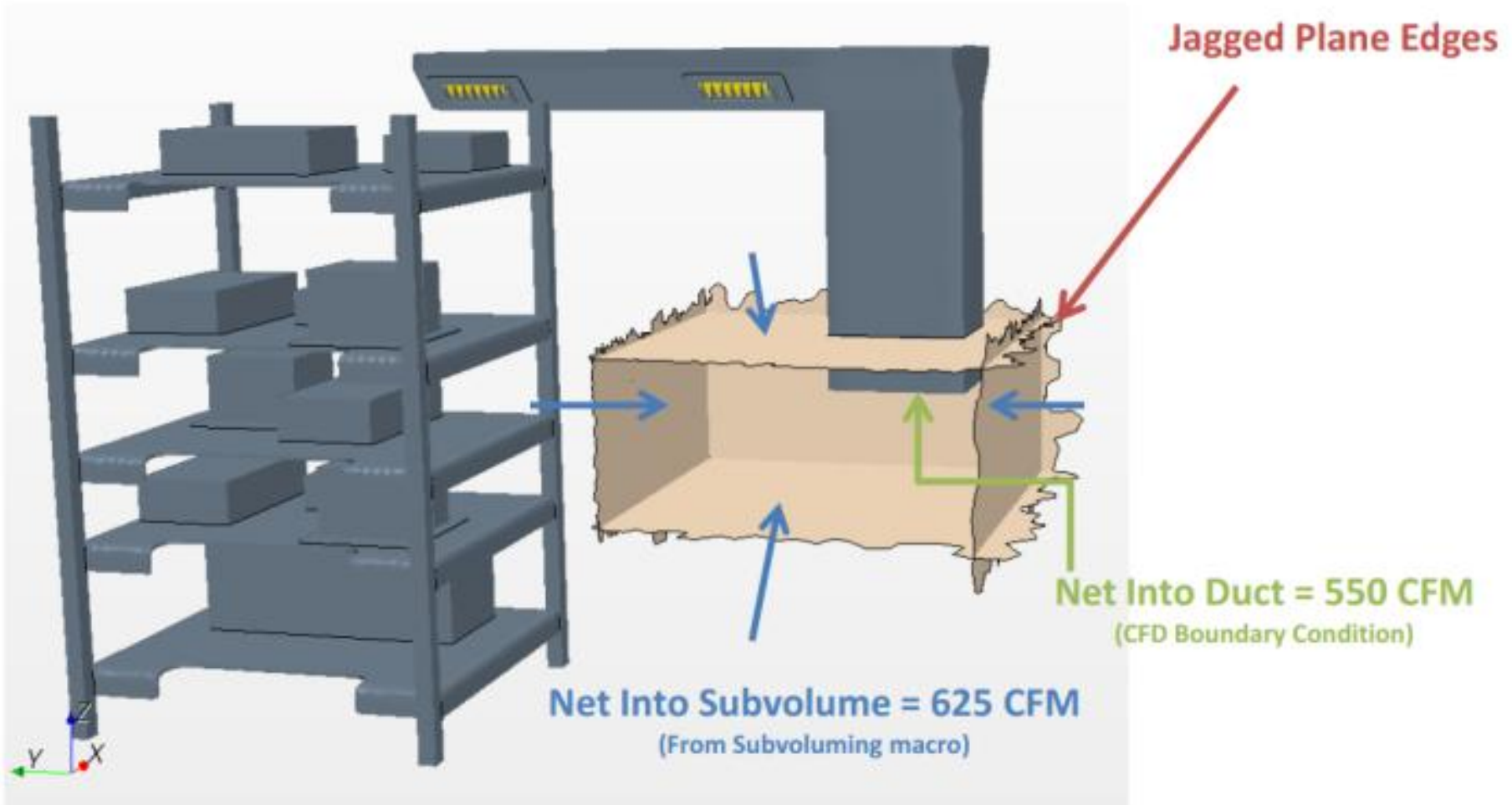
Equipment Rack



- Volumes cut complete across the shelves.
- Multiple equipment associated with each node

TECHNOLOGY DRIVEN. WARFIGHTER FOCUSED.

Another Problem with Subvoluming



Agenda

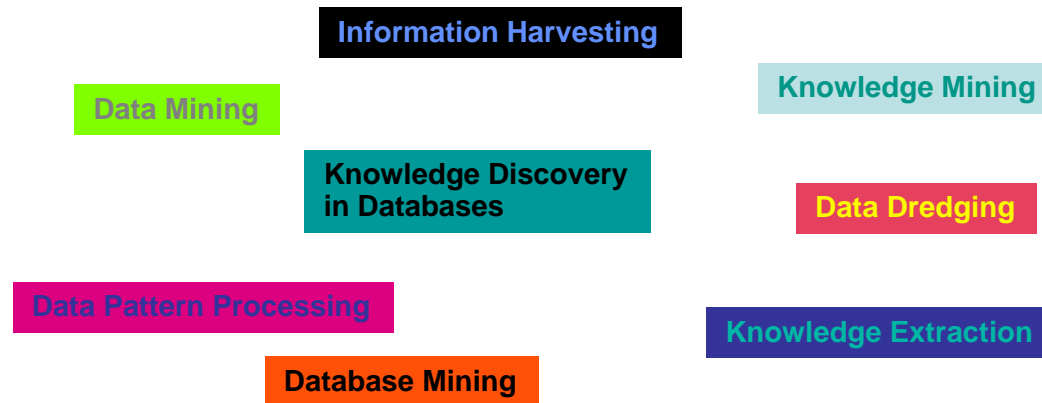


- About Rob E Smith
- Research Objectives
- Quick Introduction to Data Mining
- cfdMine Tool
- cfdMine Mahalanobis
- Validation Results
- Conclusions/ Next Steps

Data Mining



- Data mining – knowledge discovery in large sets of data
- Relatively new field in the area of computer science
- Uses/ Examples:
 - Grouping similar documents together
 - Determining consumer spending patterns (market segmentation)
 - Risk analysis and management
 - Fraud detection and detection of unusual patterns (outliers)
 - DNA and bio-data analysis

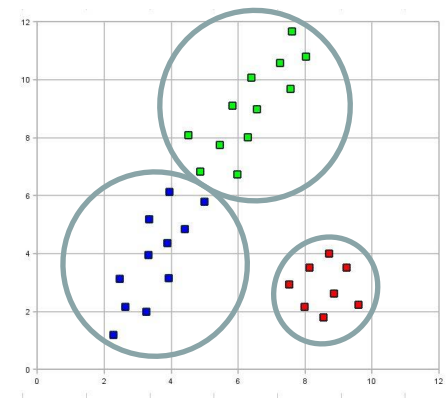


TECHNOLOGY DRIVEN. WARFIGHTER FOCUSED.

Data Mining - Clustering



- Clustering is type of data mining
- Unsupervised classification of patterns observations data items or feature vectors into groups or clusters - Jain [9]
- Clusters should be externally isolated and internally cohesive, implying homogeneity within clusters and heterogeneity between clusters - Cormack [10]
- Examples of use
 - pattern recognition
 - image processing
 - market research
 - document classification for web searching



Factors Involving Data Mining



- Selection of features of interest
 - Quantitative features:
 - Continuous values (e.g. weight).
 - Discrete values (e.g. number of computers)
 - Interval values (e.g. the duration of an event)
 - Qualitative features:
 - Nominal or unordered (e.g. color)
 - Ordinal (e.g. qualitative evaluations of temperature “cool” or “hot”)
- Preprocessing method
 - Z-normalization
 - Outlier rejection
- Selection of objective (goal) function
 - Maximize separation
 - Minimize within cluster similarity

Distance Metrics



Distances are normally used to measure similarity or dissimilarity between two data objects

- Equation 1: Euclidean Distance

$$d_e = \left(\sum_{k=1}^d (x_{i,k} - x_{j,k})^2 \right)^{\frac{1}{2}}$$

- Equation 2: Minkowski Distance

$$d_m = \left(\sum_{k=1}^d (x_{i,k} - x_{j,k})^p \right)^{\frac{1}{p}}$$

- Equation 3: Cosine Similarity (dot product)

$$d_e = \left(\sum_{k=1}^d x_{i,k} \bullet x_{j,k} \right)$$

- Equation 4: Manhattan Distance

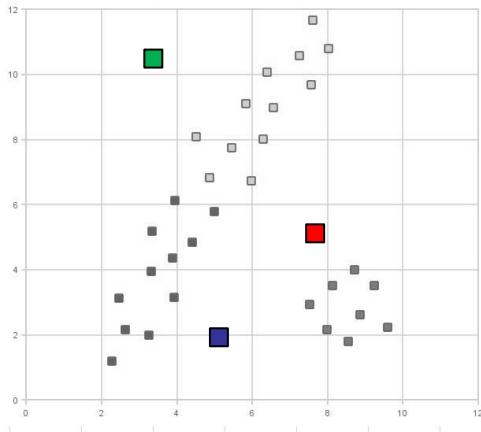
$$d_m = \sum_{k=1}^d (x_{i,k} - x_{j,k})$$

K-Means Basic Algorithm

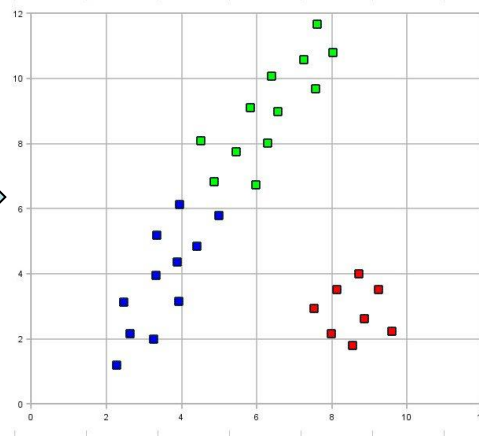


Objective is to minimize distance to cluster centers based on a Euclidean distance.

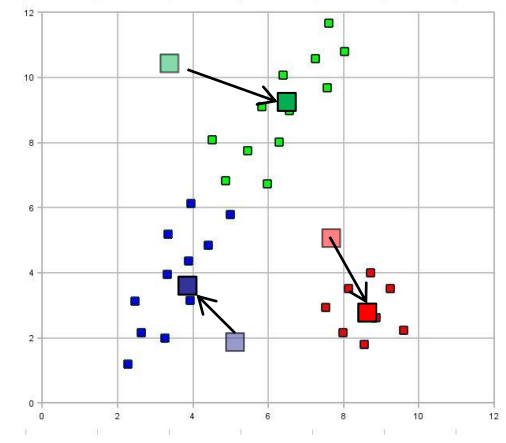
Pick Random Starting Points



Assign Data to Closest Distance Centroid

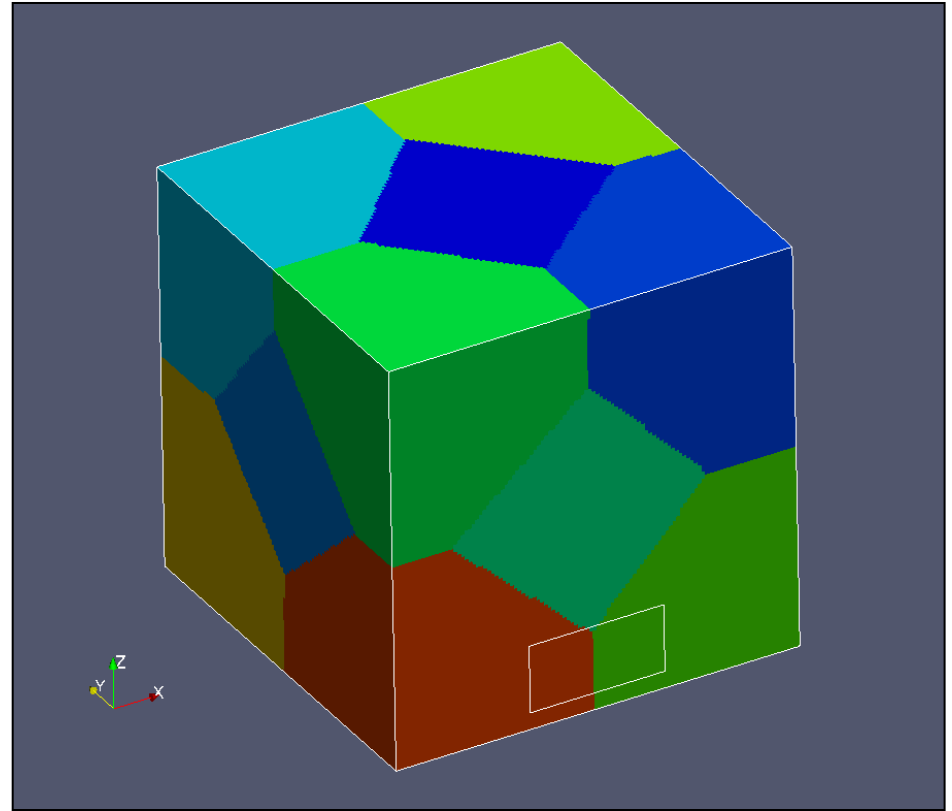
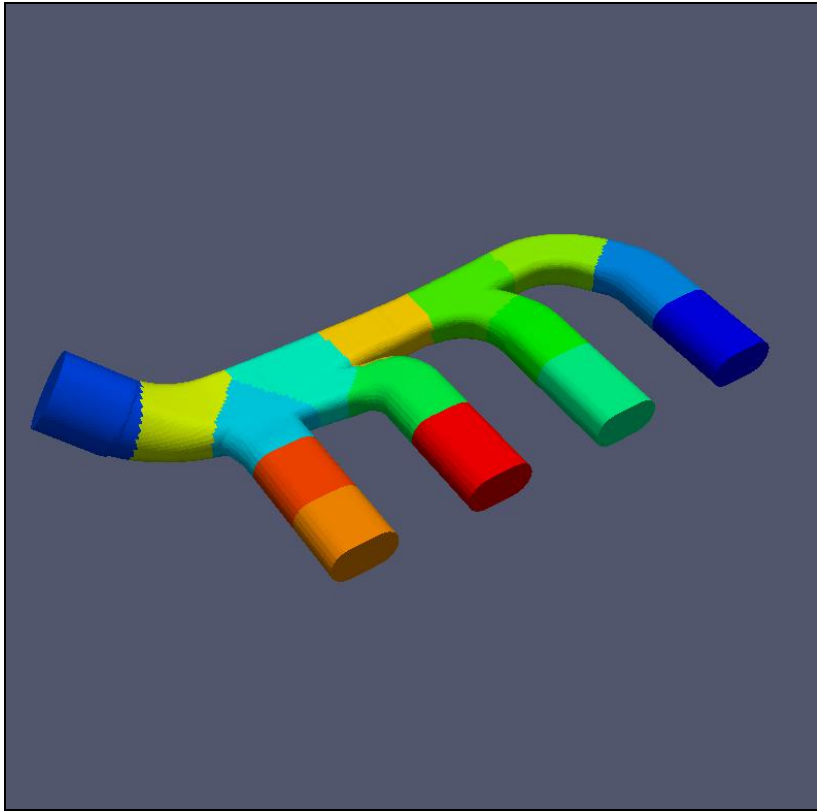


Compute a New Mean Centroid



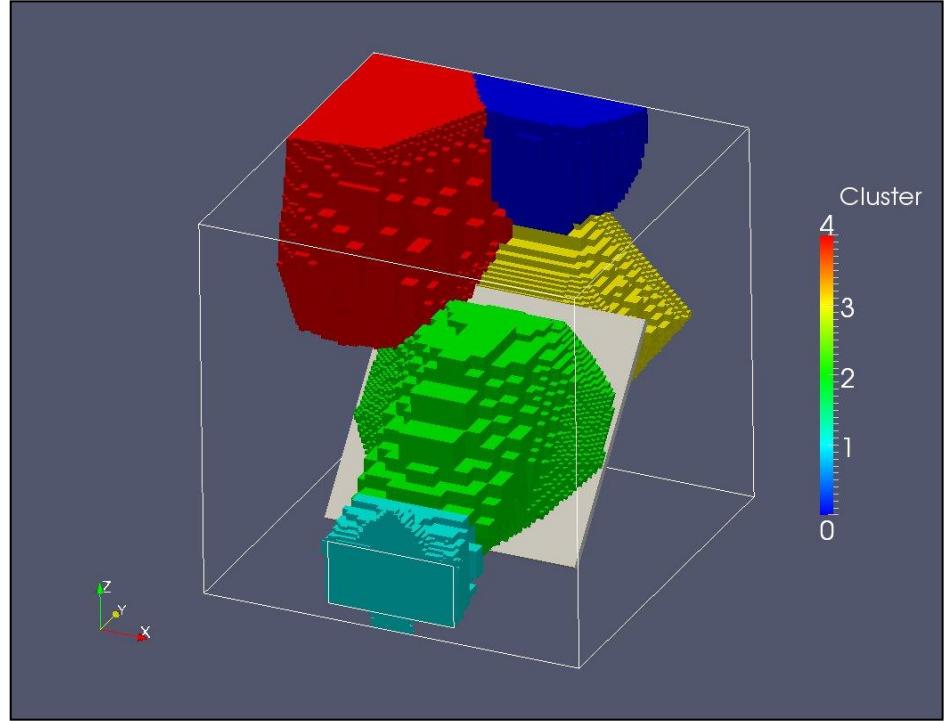
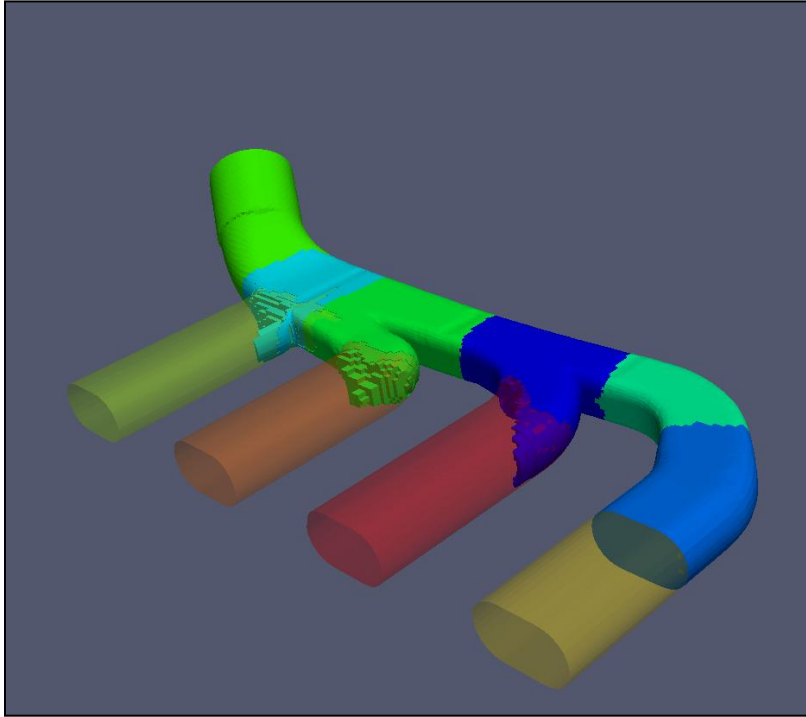
TECHNOLOGY DRIVEN. WARFIGHTER FOCUSED.

K-Means With no Feature Weighting (Only x,y,z)



Purely based on x,y,z as features, k-means produces equal "inertia" clusters about each centroid

K-Means With Weighting and Multiple Variables



- Produces equal “inertia” clusters about the weighted multivariate feature dimensions.
- Weight x, y, z distance much higher to keep clusters compact.

TECHNOLOGY DRIVEN. WARFIGHTER FOCUSED.


Agenda



- About Rob E Smith
- Research Objectives
- Quick Introduction to Data Mining
- cfdMine Tool
- cfdMine Mahalanobis
- Validation Results
- Conclusions/ Next Steps

cfdBine GUI





Ensignt Gold .geo BINARY: Case100000

Fluid Domain Part: Region 1 Load Ensignt

Clusters

15

Retries: 15

ConvergenceCrit: 0.0100

Relaxation: 1.00

MaxIter: 300

Temp Wt: 0.20

Geom Dist Wt: 0.70

Veloc Dir Wt: 0.05

Veloc Mag Wt: 0.05

Specify Fixed Clusters

2

From MuSES

From CFD

	Label	X	Y	Z	BBox	TempWt	GeoWt	VDirWt	VMagWt
1	outlet	0.1524	0.3048	0.2667	<input type="checkbox"/>	.1	1.2	0.05	0.05
2	inlet	0.1524	-1.0842...	0.0381	<input type="checkbox"/>	.1	1.2	0.05	0.05

Calculation Status

K-Means

Agglom

Calc Flux

```

K-Means Try 15 Iter 37 TotalDist 43986.7 DistChange 0.1
K-Means Try 15 Iter 38 TotalDist 43986.3 DistChange 0.4
K-Means Try 15 Iter 39 TotalDist 43986.3 DistChange 0.1
K-Means Try 15 Iter 40 TotalDist 43986.1 DistChange 0.2
K-Means Try 15 Iter 41 TotalDist 43986.1 DistChange 0.0
K-Means Try 15 Iter 42 TotalDist 43986.2 DistChange 0.1
K-Means Try 15 Iter 43 TotalDist 43986.2 DistChange 0.0
Finished Processing!
          
```

^

|

v

STOP

Write out Ensignt file for viewing

Setup in MuSES TDF file

Write Ensignt

Setup MuSES

FINISHED

cfdBine Design



- Commercial-quality C++ tool
 - QT for user interface
 - Eigen stl library for Matrices
 - Kd-tree for matching elements
 - Modified dated Ensignht reader for polyhedral mesh
 - 5000+ lines of code
- K-Means Implemented with Euclidean Distance
- Weighted features of interest
 - Temperature
 - X,Y,Z distances
 - Velocity magnitude
 - Velocity direction (as dot product)
- Preprocessing method
 - Z-normalization used

Ensignht Gold .geo BINARY: Case100000

TARDEC Fluid Domain Part: Region 1 Load Ensignht

Clusters: 15

Retries: 15 ConvergenceCrit: 0.0100 Relaxation: 1.00 MaxIter: 300

Temp Wt: 0.20 Geom Dist Wt: 0.70 Veloc Dir Wt: 0.05 Veloc Mag Wt: 0.05

Specify Fixed Clusters: 2

Label	X	Y	Z	BBox	TempWt	GeoWt	VDirWt	VMagWt
1 outlet	0.1524	0.3048	0.2667		.1	1.2	0.05	0.05
2 inlet	0.1524	-1.0842...	0.0381		.1	1.2	0.05	0.05

From MuSES From CFD

Calculation Status

K-Means Agglom Calc Flux

```

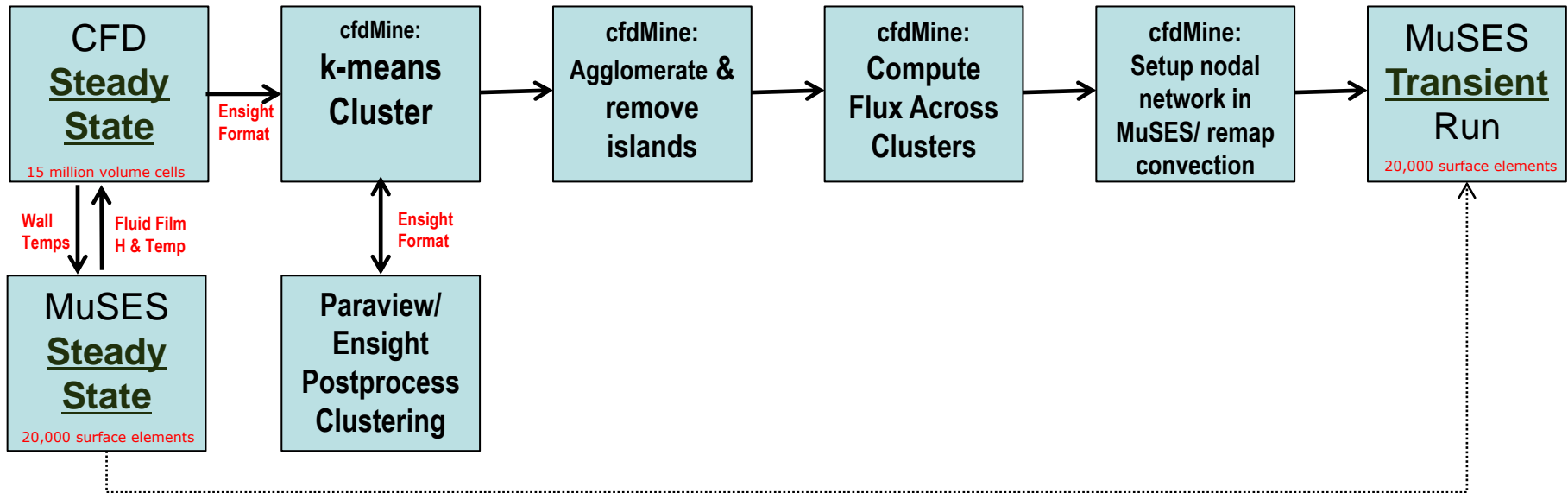
K-Means Try 15 Iter 37 TotalDist 43986.7 DistChange 0.1
K-Means Try 15 Iter 38 TotalDist 43986.3 DistChange 0.4
K-Means Try 15 Iter 39 TotalDist 43986.3 DistChange 0.1
K-Means Try 15 Iter 40 TotalDist 43986.1 DistChange 0.2
K-Means Try 15 Iter 41 TotalDist 43986.1 DistChange 0.0
K-Means Try 15 Iter 42 TotalDist 43986.2 DistChange 0.1
K-Means Try 15 Iter 43 TotalDist 43986.2 DistChange 0.0
Finished Processing!
  
```

Write out Ensignht file for viewing Write Ensignht FINISHED

Setup in MuSES TDF file Setup MuSES

TECHNOLOGY DRIVEN. WARFIGHTER FOCUSED.

Basic cfdMine Process



Note: The nodal network is set up in MuSES using a thermal link from every element to each cluster fluid node. (A first, but not ideal)

Which Knob Does What



Ensignt Gold .geo BINARY: Case100000

TARDEC Fluid Domain Part: Region 1 Load Ensignt

Clusters: 15

Retries: 15 ConvergenceCrit: 0.0100 Relaxation: 1.00 MaxIter: 300

Temp Wt: 0.20 Geom Dist Wt: 0.70 Veloc Dir Wt: 0.05 Veloc Mag Wt: 0.05

Specify Fixed Clusters: 2

From MuSES
From CFD

	Label	X	Y	Z	BBox	TempWt	GeoWt	VDirWt	VMagWt
1	outlet	0.1524	0.3048	0.2667	<input type="checkbox"/>	.1	1.2	0.05	0.05
2	inlet	0.1524	-1.0842...	0.0381	<input type="checkbox"/>	.1	1.2	0.05	0.05

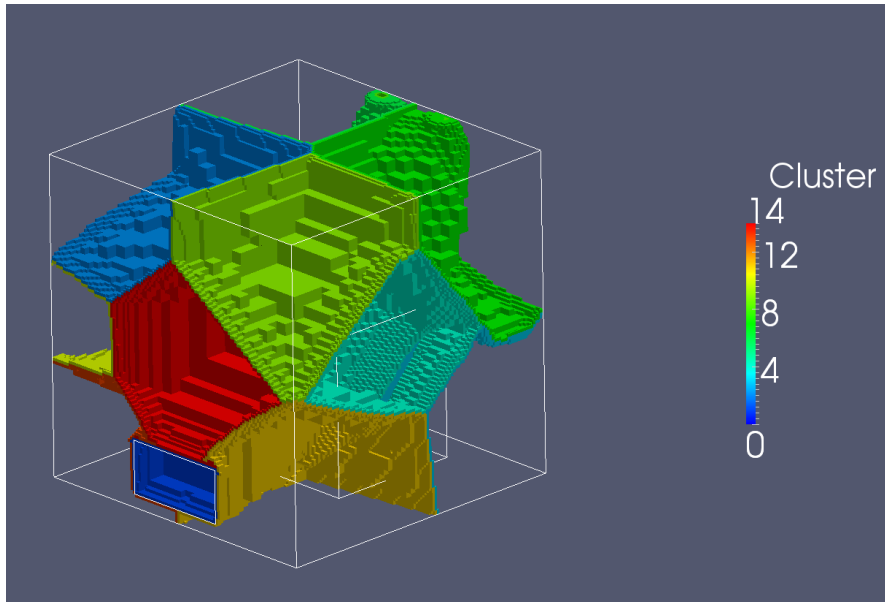
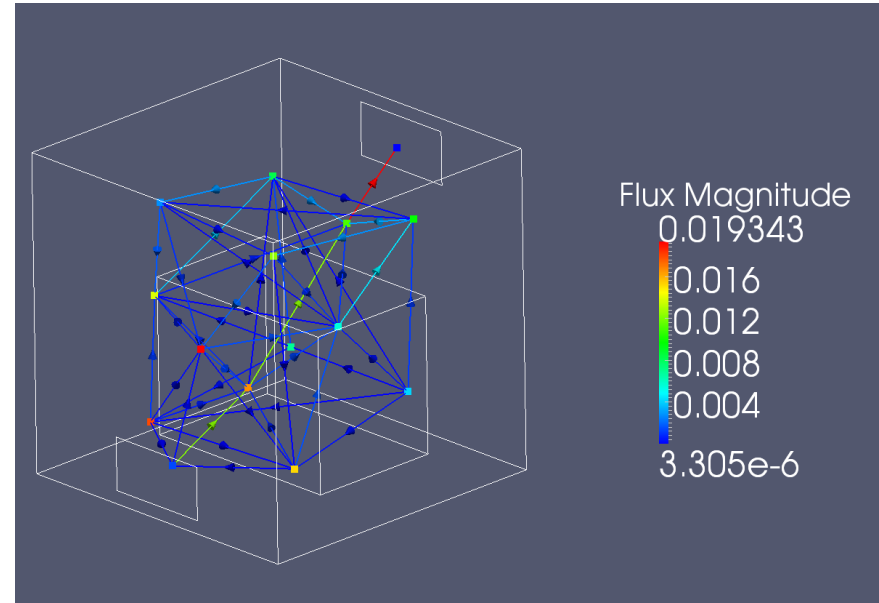
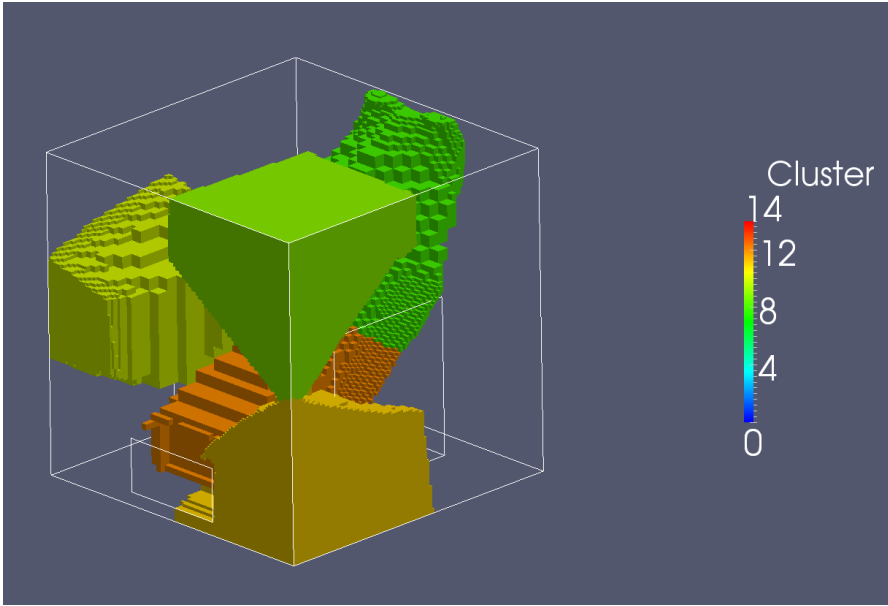
Calculation Status

K-Means
Agglom
Calc Flux

K-Means Try 15 Iter 37 TotalDist 43986.7 DistChange 0.1
 K-Means Try 15 Iter 38 TotalDist 43986.3 DistChange 0.4
 K-Means Try 15 Iter 39 TotalDist 43986.3 DistChange 0.1
 K-Means Try 15 Iter 40 TotalDist 43986.1 DistChange 0.2
 K-Means Try 15 Iter 41 TotalDist 43986.1 DistChange 0.0
 K-Means Try 15 Iter 42 TotalDist 43986.2 DistChange 0.1
 K-Means Try 15 Iter 43 TotalDist 43986.2 DistChange 0.0
 Finished Processing!

Write out Ensignt file for viewing Write Ensignt
 Setup in MuSES TDF file Setup MuSES FINISHED

Visualizing Clusters – Case 1

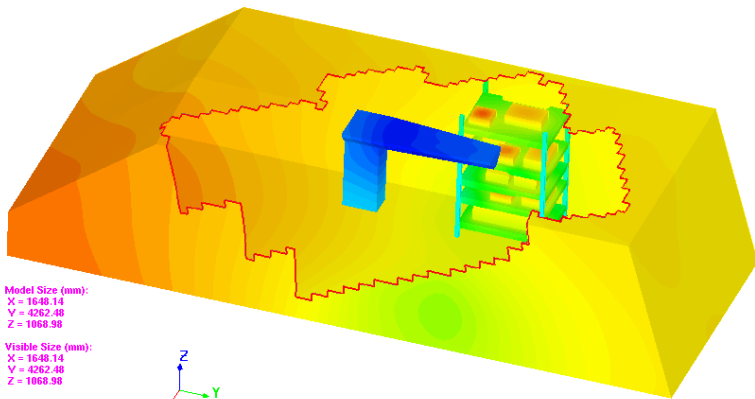
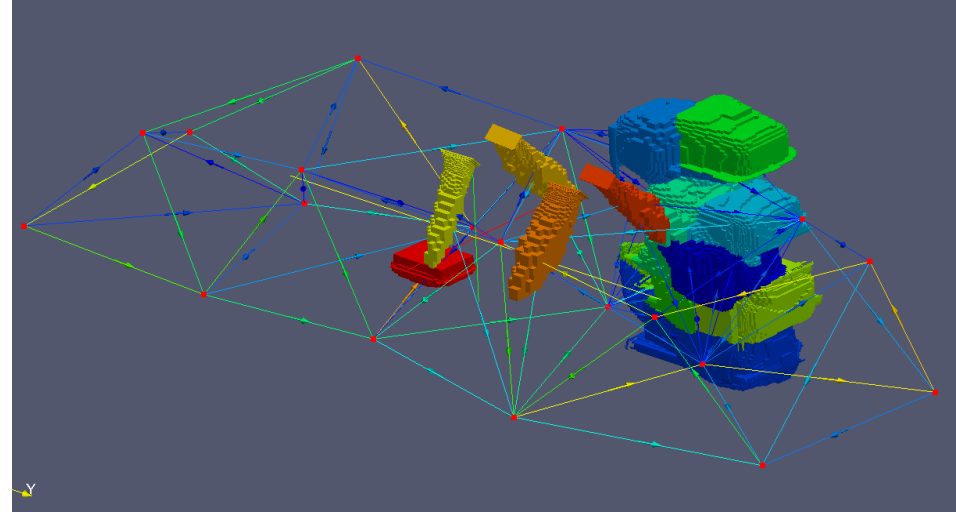
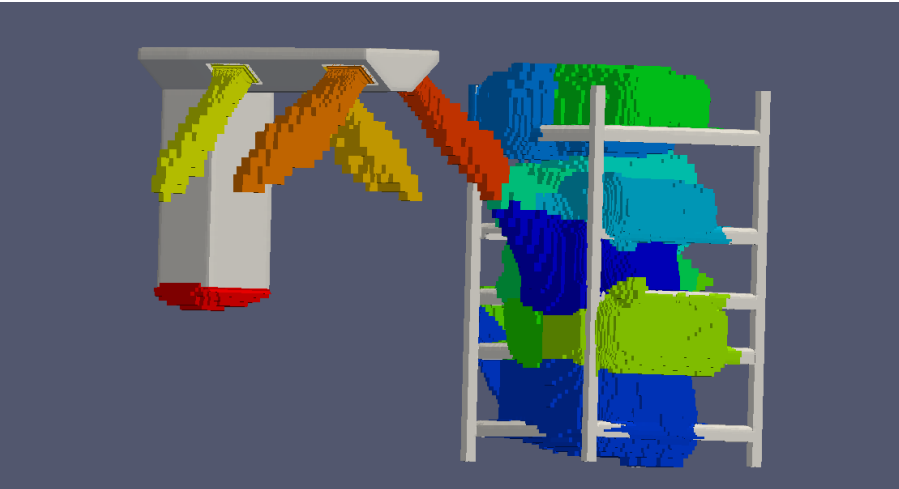


TECHNOLOGY DRIVEN. WARFIGHTER FOCUSED.

Visualizing Clusters – Generic Vehicle Interior



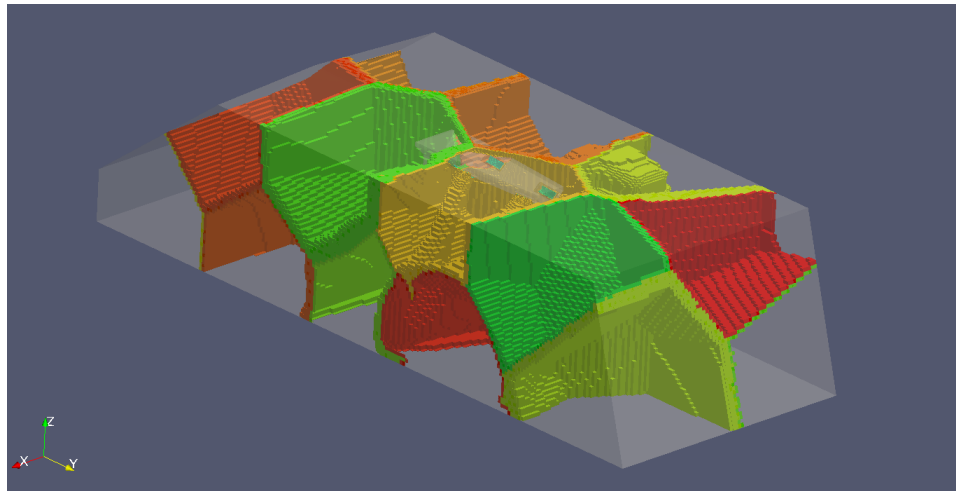
UNCLASSIFIED



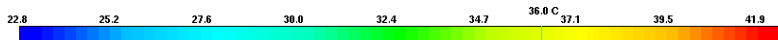
Model Size (mm):
X = 1648.14
Y = 4262.48
Z = 1068.98

Visible Size (mm):
X = 1648.14
Y = 4262.48
Z = 1068.98

Visible Counts:
Parts = 15
Elements = 48255

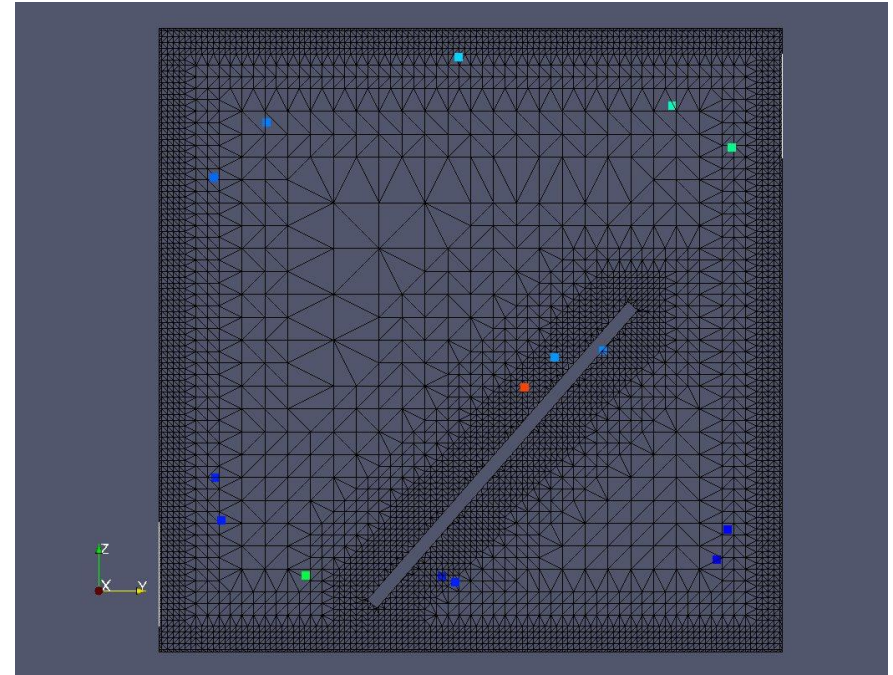
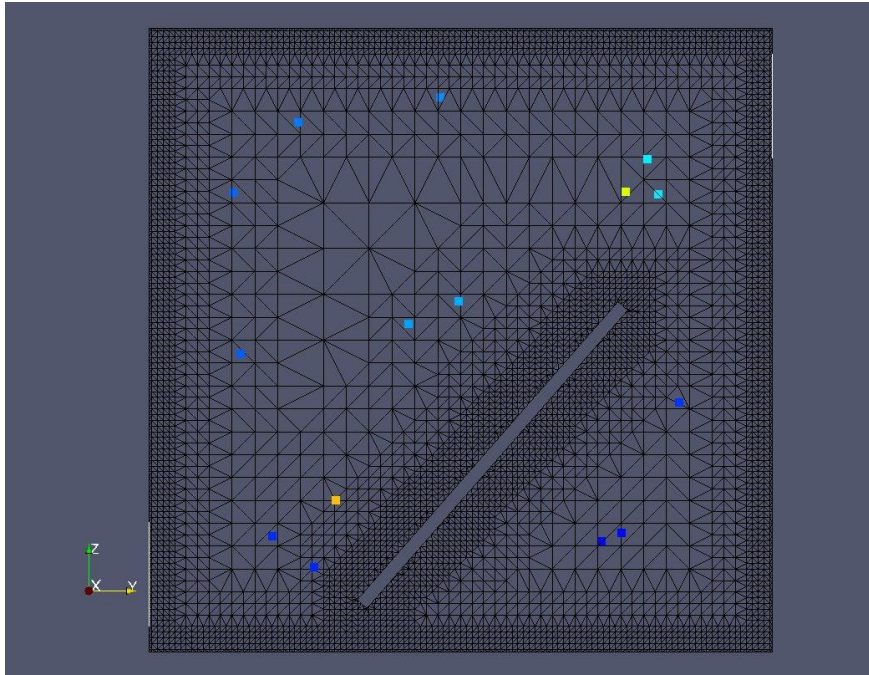


TECHNOLOGY DRIVEN. WARFIGHTER FOCUSED.



UNCLASSIFIED

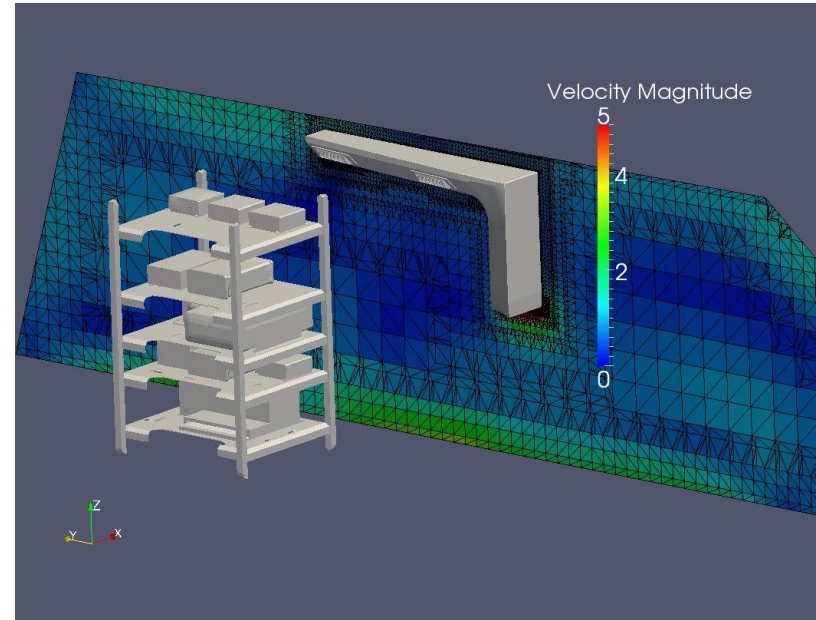
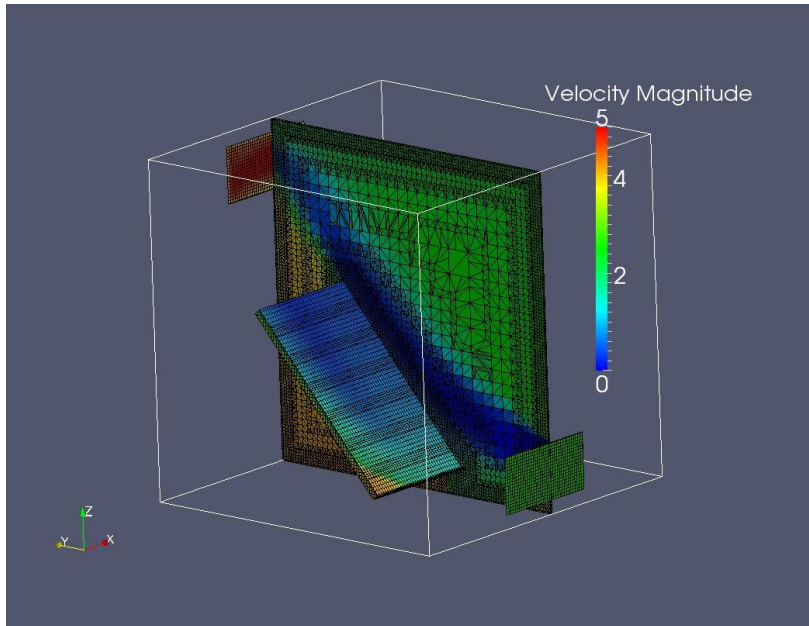
Importance of Volume Weighting



K-Means Volume Weighted Iterative Function

$$\bar{m}_l = \frac{\sum_{x \in S_l} v_x}{n_l} \sum_{x \in S_l} \vec{x} \cdot v_x$$

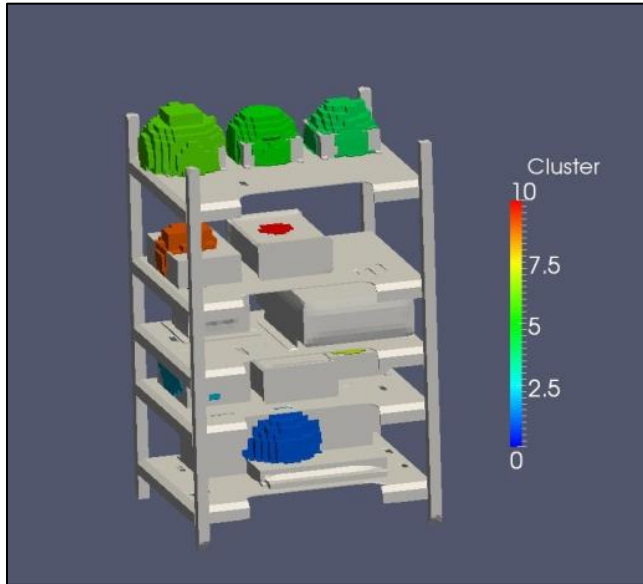
Reason To Weight – Typical Meshes



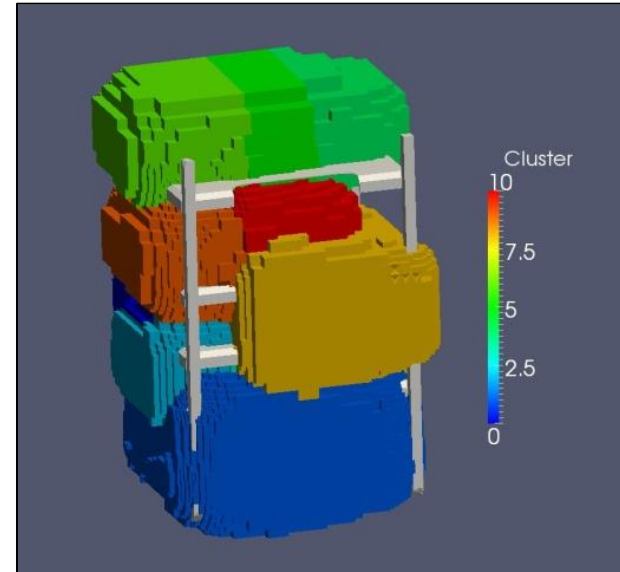
Fixing Clusters Around Equipment



Clustering About the Centroid



Clustering About the Bounding Box

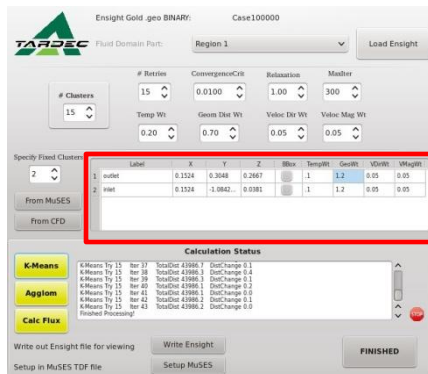


$$x_{bbox_min} = (x < x_{min}) ? x_{min} : (x > x_{max}) ? x_{max} : x$$

$$y_{bbox_min} = (y < y_{min}) ? y_{min} : (y > y_{max}) ? y_{max} : y$$

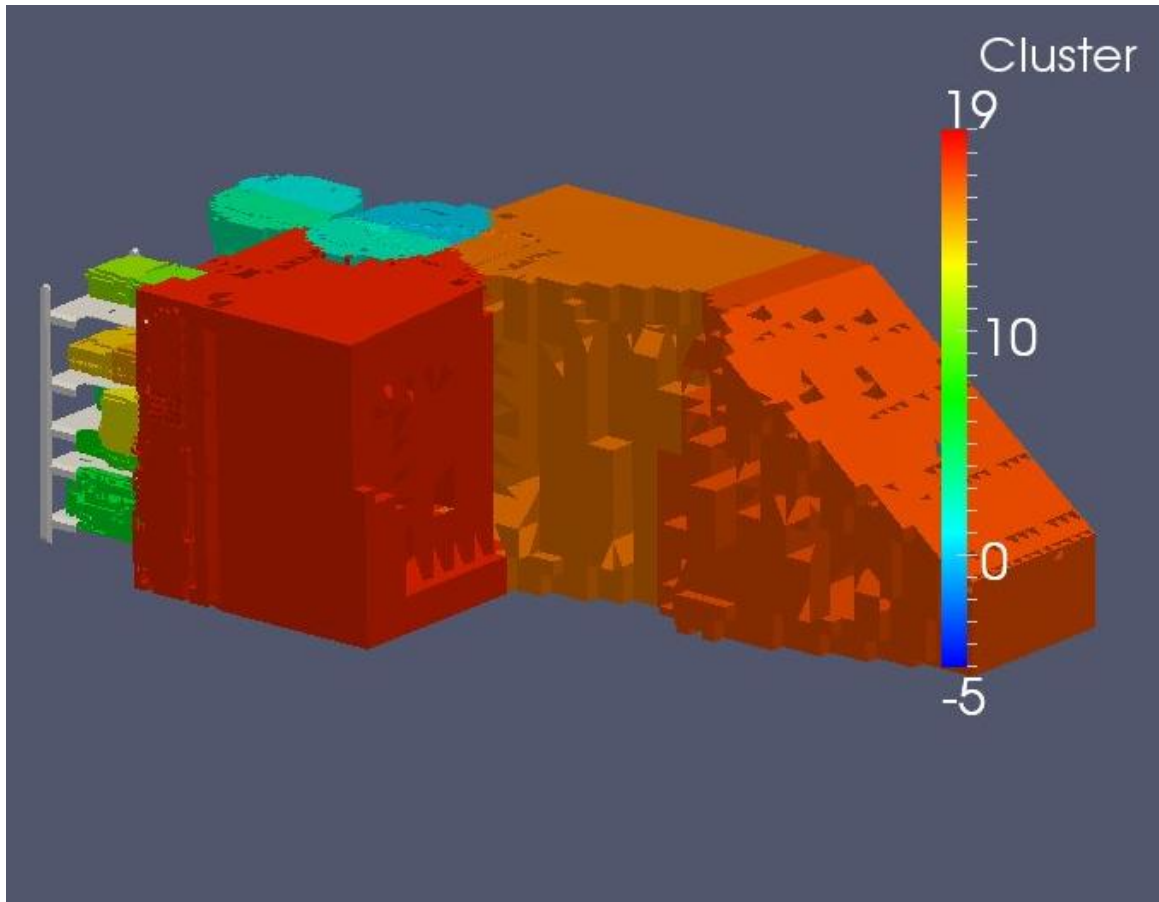
$$z_{bbox_min} = (z < z_{min}) ? z_{min} : (z > z_{max}) ? z_{max} : z$$

$$dist_{min} = \sqrt{(x - x_{bbox_min})^2} + \sqrt{(y - y_{bbox_min})^2} + \sqrt{(z - z_{bbox_min})^2}$$



Clustering around the bounding box is quick and efficient.

Cluster Size Around Equipment



Enight Gold .geo BINARY: Case100000

TARDEC Fluid Domain Part: Region 1 Load Enight

Clusters: 15

Retries: 15

ConvergenceCrit: 0.0100

Relaxation: 1.00

MaxIter: 300

Temp Wt: 0.20

Geom Dist Wt: 0.70

Veloc Dir Wt: 0.05

Veloc Mag Wt: 0.05

Specify Fixed Clusters

Label	X	Y	Z	BBx	BBz	TempWt	GeoWt	VDirWt	VMagWt
1 outlet	0.1524	0.3048	0.2667			.1	1.2	0.05	0.05
2 inlet	0.1524	-1.0842...	0.0381			.1	1.2	0.05	0.05

From MuSES

From CFD

Calculation Status

K-Means

Agglom

Calc Flux

Write out Enight file for viewing Write Enight

Setup in MuSES TDF file Setup MuSES

FINISHED

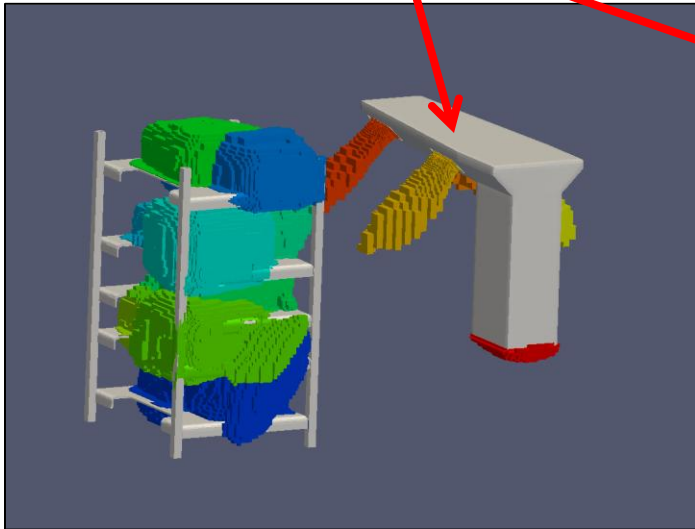
Larger weights allow the fixed location equipment-clusters to be smaller

UNCLASSIFIED

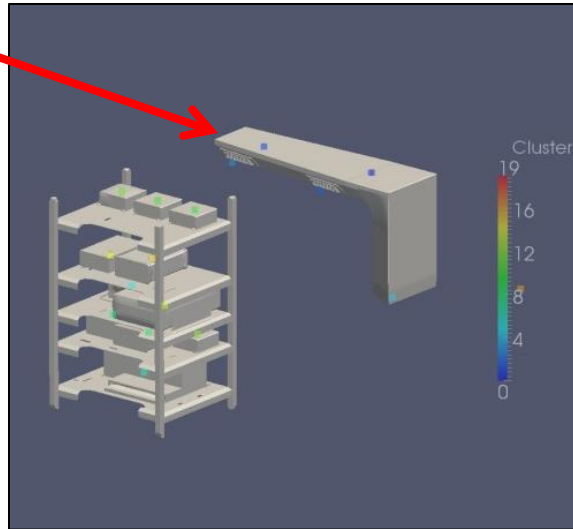
Fixing Clusters Around Inflow & Outflows



Important to Place Clusters at Inflow/ Outflows

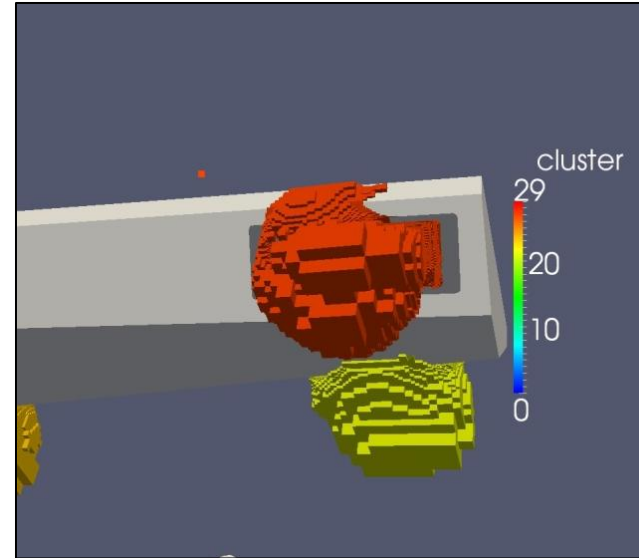


Cluster center locations



- Required to enforce flow network conservation of volume flow
- Usually an area of interest

Bifurcation Problem at Inflow/Outflow

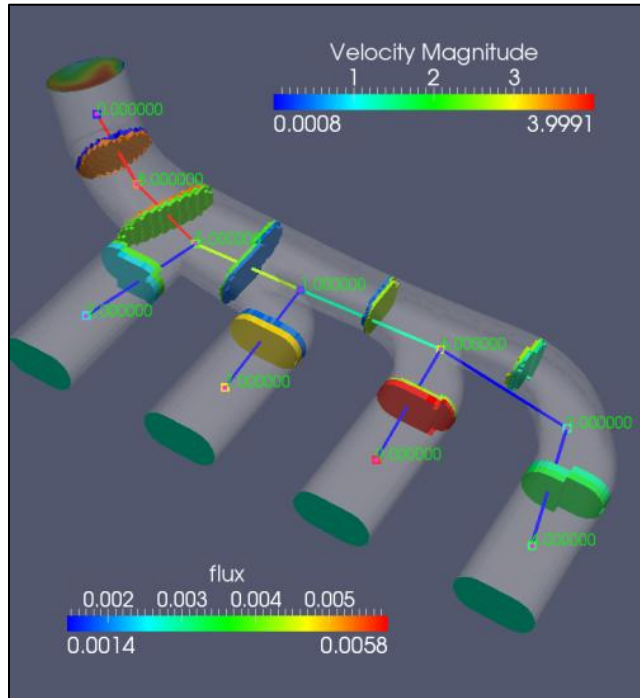


- Usually better to use centroid distance (as opposed to bounding box) from inflow/ outflow
- Added feature to specify cluster temperature, velocity at fixed nodes, to further avoid this

Validating Flux at The Interface Boundaries



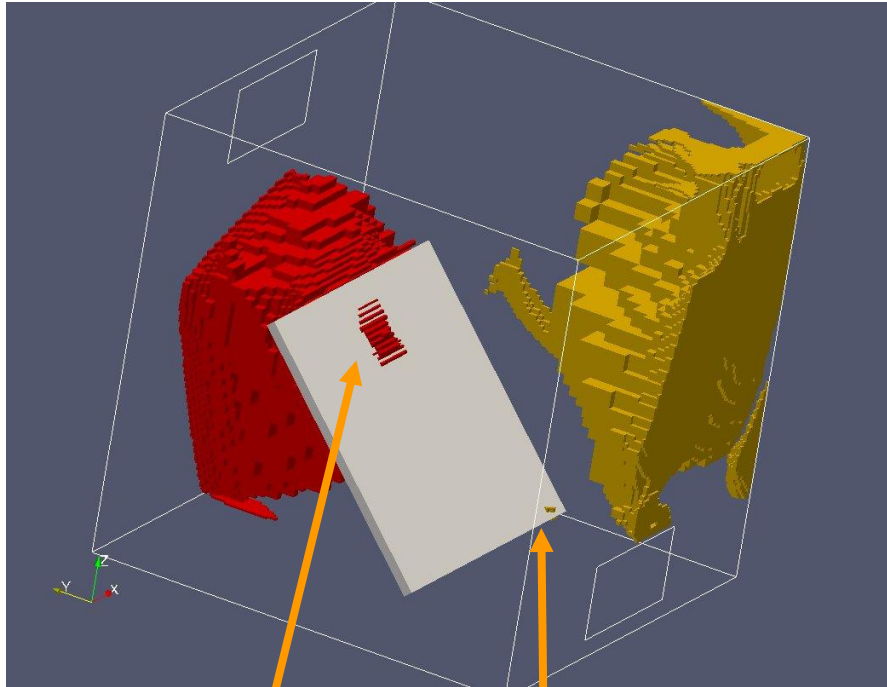
$$interface\ flux = \sum_1^{N_{faces}} a \cdot \frac{\vec{n} \cdot \vec{v}_1 + \vec{n} \cdot \vec{v}_2}{2}$$



From Cluster	To Cluster	cfDMine m ³ /s	StarCCM+ m ³ /s	Error
8	0	0.005785	0.005798	0.23%
1	5	0.004345	0.004353	0.19%
6	1	0.002899	0.002908	0.33%
7	1	0.001445	0.001447	0.12%
2	5	0.001442	0.001445	0.19%
4	3	0.001447	0.001448	0.05%
3	6	0.001447	0.001448	0.05%
5	8	0.005788	0.005798	0.17%
9	6	0.001446	0.001445	-0.08%

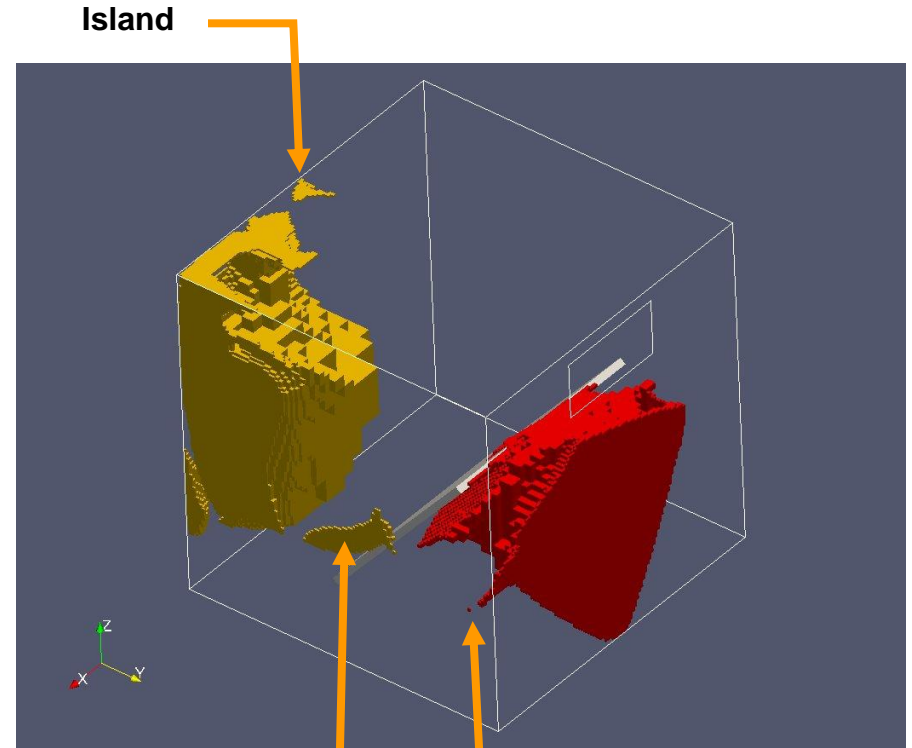
This case is fairly idealized: Planes are perpendicular to flow, fine mesh, and no major temperature gradients.

Agglomeration – Dealing With Islands



Island

Island



Island

Island

Island

Enlight Gold geo BINARY: Case1000000

TARDEC Fluid Domain Part: Region.1 Load Enight

Clusters: 15

Iterates: 15 ConvergenceCrt: 0.0100 Relaxation: 1.00 MaxIter: 300

Temp Wt: 0.20 Geom Ekt Wt: 0.70 Veloc Ekt Wt: 0.05 Veloc Mag Wt: 0.05

Speedy Fixed Clusters	Label	X	Y	Z	IBloc	TempWt	GeomWt	VelocWt	VMagWt
1	outlet	0.1524	0.3048	0.2667	1	1.2	0.05	0.05	
2	inlet	0.1524	-1.0842	0.0395	1	1.2	0.05	0.05	

Calculation Status

K-Means	K-Means Try	Iter	TotalDx	13995.1	DistChange	0.1
1.0	Try 15	Iter 37	TotalDx	43995.3	DistChange	0.4
1.0	Try 25	Iter 38	TotalDx	43995.3	DistChange	0.4
1.0	Try 35	Iter 39	TotalDx	43995.3	DistChange	0.1
1.0	Try 45	Iter 40	TotalDx	43995.1	DistChange	0.2
1.0	Try 55	Iter 41	TotalDx	43995.1	DistChange	0.0
1.0	Try 65	Iter 42	TotalDx	43995.2	DistChange	0.1
1.0	Try 75	Iter 43	TotalDx	43995.2	DistChange	0.0

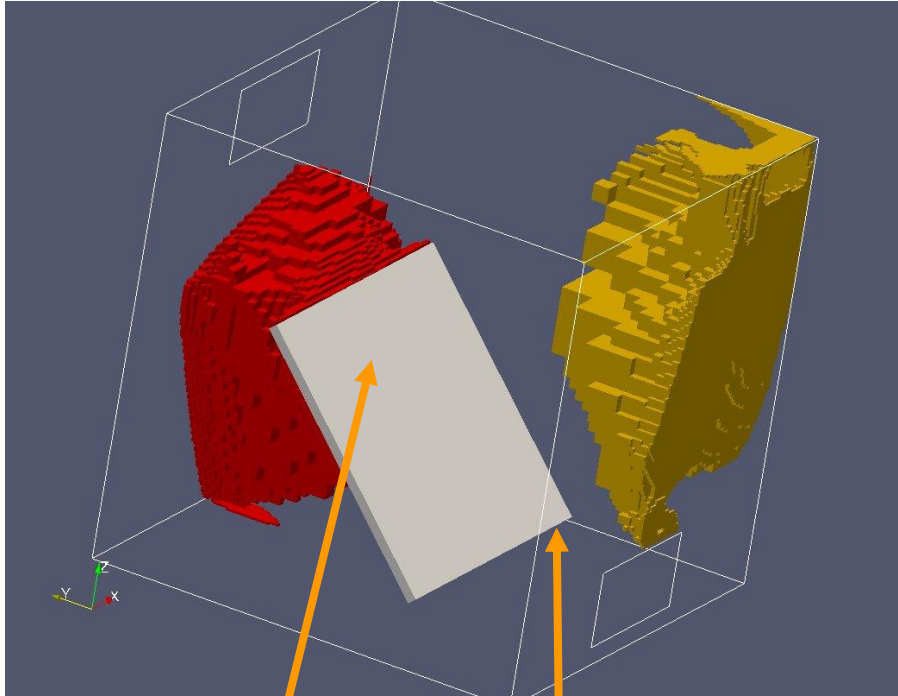
Agglomeration

Write out Enight file for viewing Write Enight FINISHED

Setup in MUSES TDF file Setup MUSES

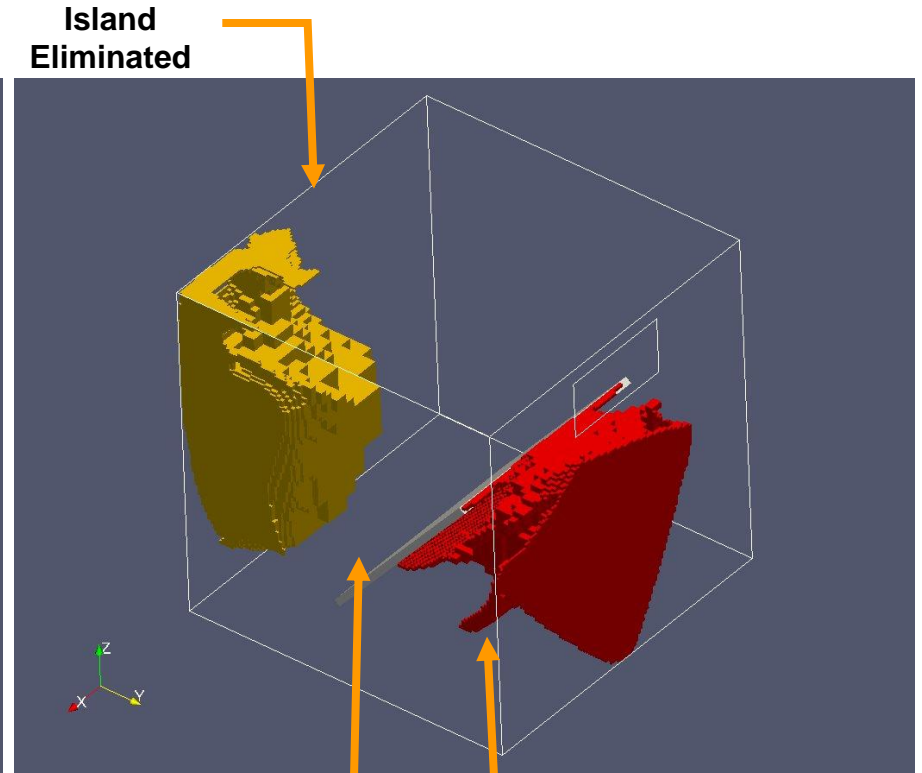
TECHNOLOGY DRIVEN. WARFIGHTER FOCUSED.

Agglomeration – Dealing With Islands



Island
Eliminated

Island
Eliminated

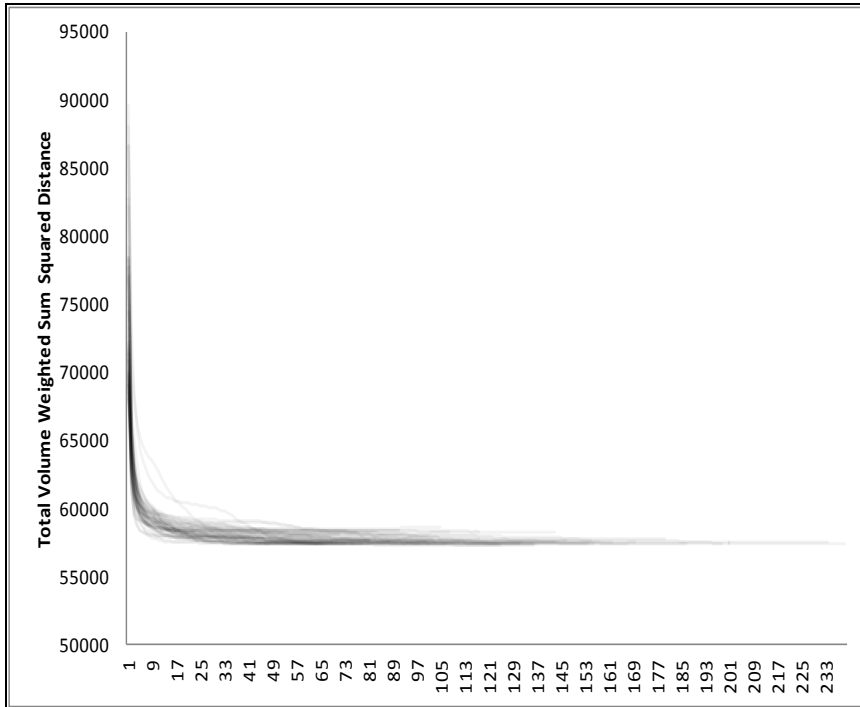


Island
Eliminated

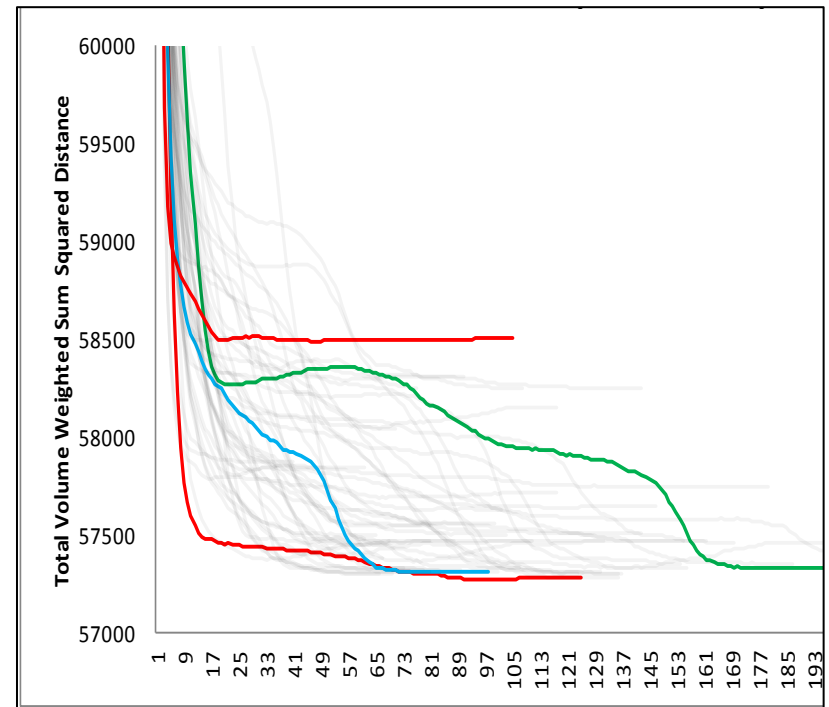
Island
Eliminated

Island
Eliminated

Convergence Plots



Typical k-means convergence plot for 50 random start trials



"zoomed-in" plot of 50 random-start k-means trials

Agenda



- About Rob E Smith
- Research Objectives
- Quick Introduction to Data Mining
- cfdMine Tool
- cfdMine Mahalanobis
- Validation Results
- Conclusions/ Next Steps

cfdBine Mahalanobis



File About

Ensignt Gold .geo BINARY: subvolume_validation_final_trans-fine2@6.00e+010

Fluid Domain Part: F_Cabin Load Ensignt

Clusters: 35

Retries: 10 ConvergenceCrit: 0.0050 Init Euc Iters: 10 Max Iter: 40 Covar Update: 10

Temp Wt: 1.00 Geom Dist Wt: 2.00 Veloc Dir Wt: 1.00 Veloc Mag Wt: 1.00

Euclidean Shape
Mahalanobis

Specify Fixed Clusters: 15

From MuSES
From CFD

	Label	X	Y	Z	BBox	TempWt	GeoWt	VDirWt	VMag
1	Duct	0.019268	4.82436	2.68189	<input checked="" type="checkbox"/>	1	5	1	1
2	C_Relay_Box	-0.523846	5.33141	2.47283	<input checked="" type="checkbox"/>	1	5	1	1
3	C_RPCU	-0.547708	5.20134	2.1773	<input checked="" type="checkbox"/>	1	5	1	1
4	C_DCE_1	-0.526248	5.02561	2.82669	<input checked="" type="checkbox"/>	1	5	1	1

Calculation Status

K-Means Mahalanobis
Agglom
Calc Flux

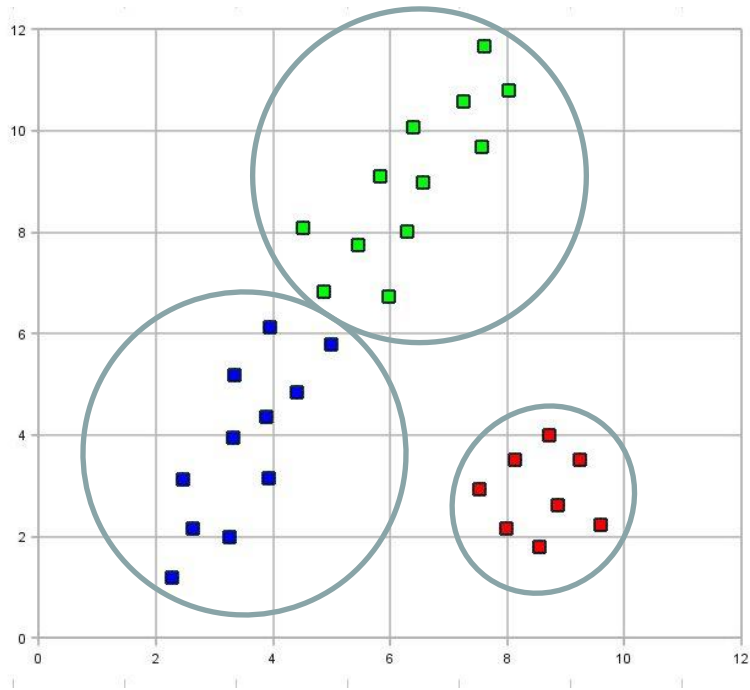
Reading Ensignt Gold geometry file...
Reading element results data...
Temperatures
Centroids
Velocities
Volumes
Populating data structures...

Write out Ensignt file for viewing Write Ensignt

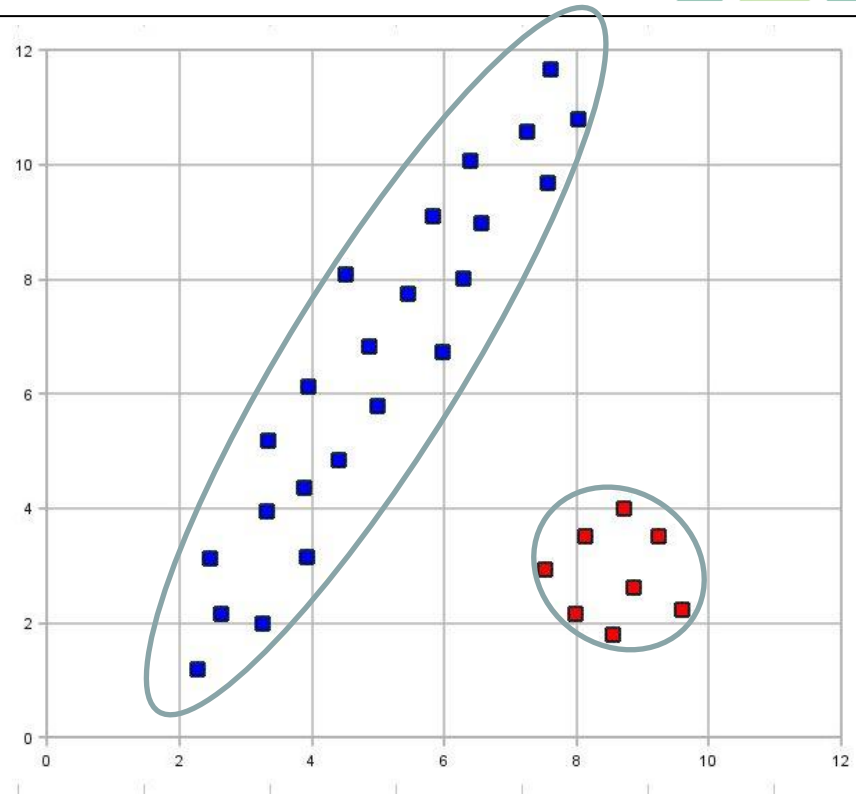
Setup in MuSES TDF file Setup MuSES

FINISHED

Mahalanobis Distance Metric



Euclidean Distance
(Standard K-Means)



Mahalanobis Distance

Euclidean distances fit clusters to hyperspheres.
Mahalanobis fits data to an oriented hyperellipsoid.

Mahalanobis Distance



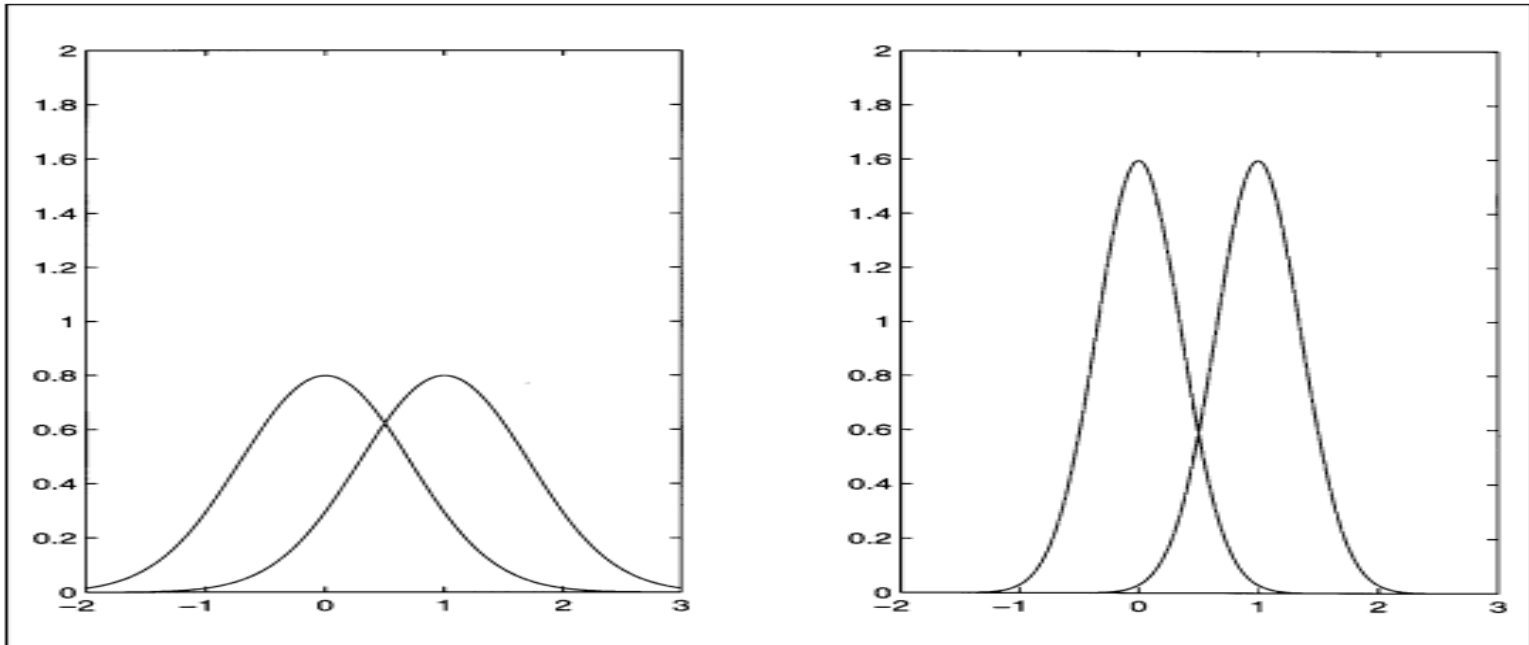
$$\Delta^2 = (x_i - \mu_k)^T \Sigma^{-1} (x_i - \mu_k) , \text{ where}$$

$$\Sigma = \begin{bmatrix} \sigma_{11} & \sigma_{12} & \sigma_{1L} \\ \sigma_{21} & \sigma_{22} & \sigma_{2L} \\ \sigma_{L1} & \sigma_{L2} & \sigma_{LL} \end{bmatrix} \quad \text{covariance matrix}$$

Note: An identity matrix results in Euclidean distance.

Mahalanobis cannot be used directly with K-means. Adds 36 unknowns and drops into non-optimal local minima even faster.

Example: Univariate Mahalanobis Distances



Euclidean distances are one in both cases, but Mahalanobis distance of $\Delta = 1$ and 2

cfdBine Mahalanobis / New Algorithm



File About

Enight Gold .geo BINARY: subvolume_validation_final_trans-fine2@6.00e+010



Fluid Domain Part:

F_Cabin

Load Enight

Clusters

35

Retries

10

ConvergenceCrit

0.0050

Init Euc Iters

10

Max Iter

40

Covar Update

10

Euclidean

Temp Wt

1.00

Geom Dist Wt

2.00

Veloc Dir Wt

1.00

Veloc Mag Wt

1.00

Shape

Mahalanobis

Specify Fixed Clusters

15

From MuSES

From CFD

	Label	X	Y	Z	BBox	TempWt	GeoWt	VDirWt	VMag
1	Duct	0.019268	4.82436	2.68189	<input checked="" type="checkbox"/>	1	5	1	1
2	C_Relay_Box	-0.523846	5.33141	2.47283	<input checked="" type="checkbox"/>	1	5	1	1
3	C_RPCU	-0.547708	5.20134	2.1773	<input checked="" type="checkbox"/>	1	5	1	1
4	C_DCE_1	-0.526248	5.02561	2.82669	<input checked="" type="checkbox"/>	1	5	1	1

Calculation Status

K-Means
Mahalanobis

Agglom

Calc Flux

Reading Enight Gold geometry file...
 Reading element results data...
 Temperatures
 Centroids
 Velocities
 Volumes
 Populating data structures...

STOP

Write out Enight file for viewing

Write Enight

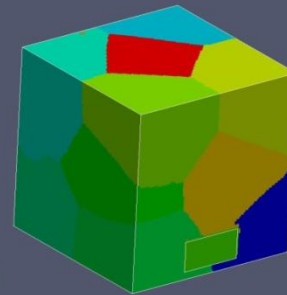
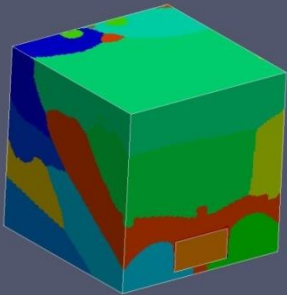
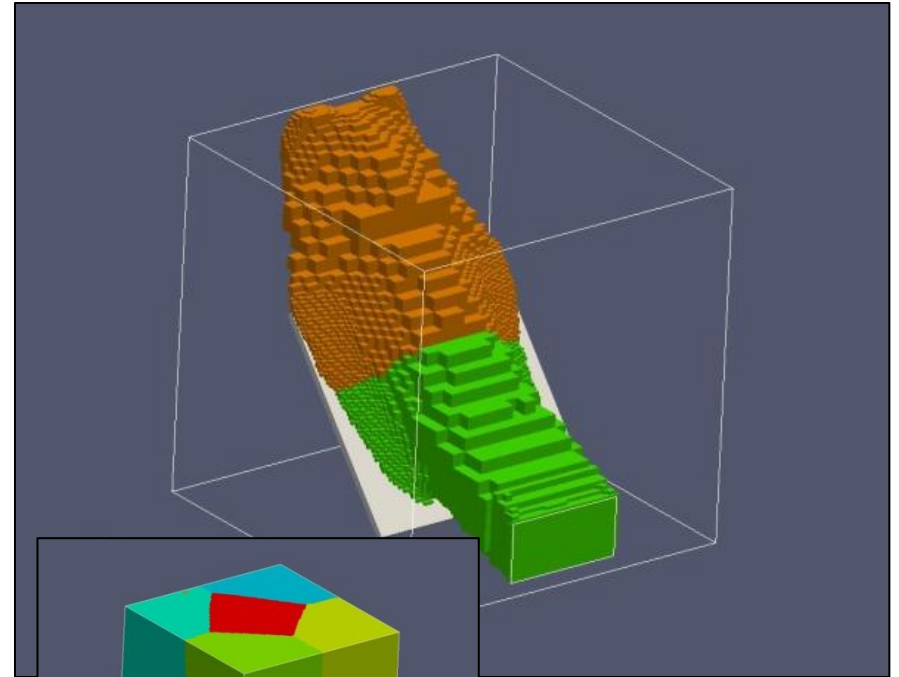
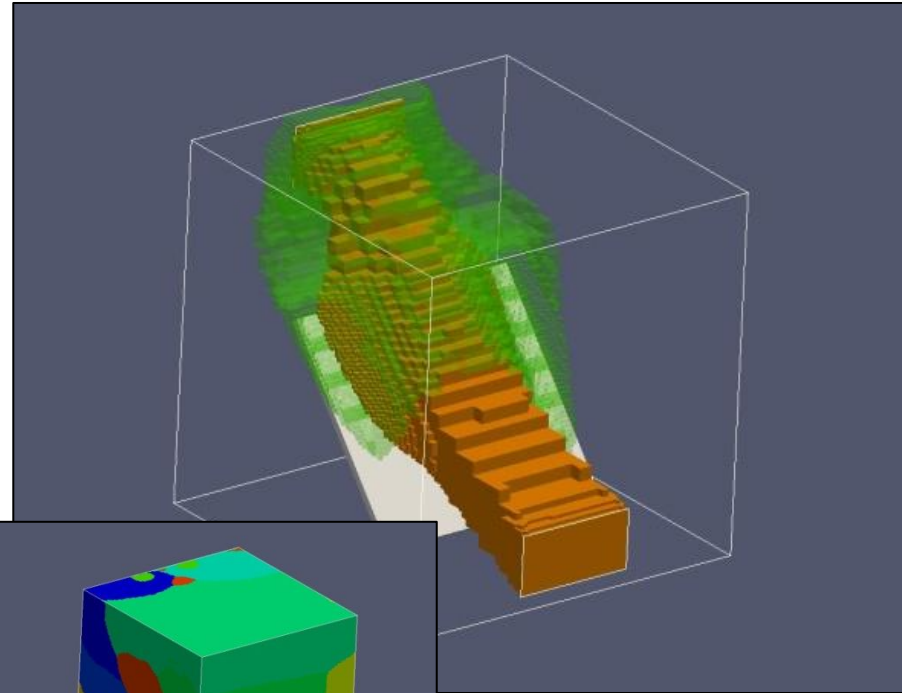
Setup in MuSES TDF file

Setup MuSES

FINISHED

- Initial "m" pure Euclidean iterations
- Blend inv covariance w/ Euclidean(identity)
- Only update covariance "n" iterations

Results – Better Data Separation



Mahalanobis clusters

15 clusters, 75% Euclidean and 15% Mahalanobis

Pure Euclidean

15 clusters, 100% Euclidean

Mahalanobis resolves entire "jet" in Case 1, inlet to outlet.

Agenda

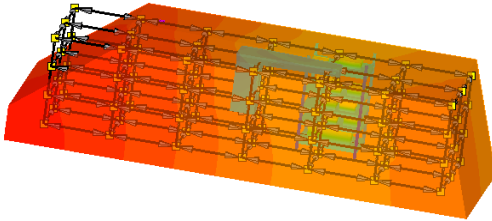


- About Rob E Smith
- Research Objectives
- Quick Introduction to Data Mining
- cfdMine Tool
- cfdMine Mahalanobis
- Validation Results
- Conclusions/ Next Steps

Validation Goal/ Method

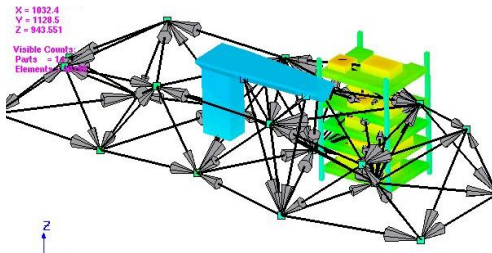


- Subvolumed transient MuSES nodal network



- Based on steady-state CFD
- Transient 30 minute run

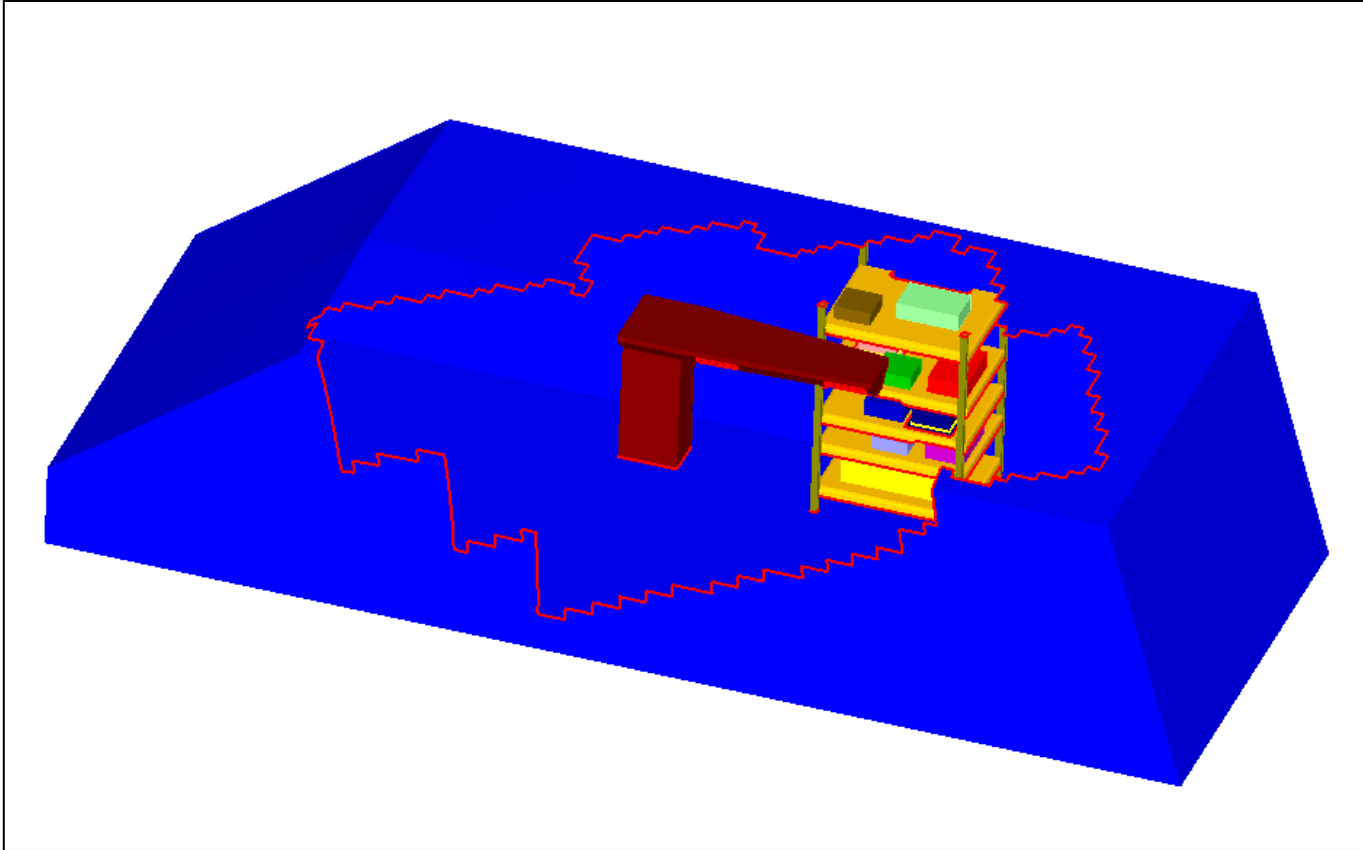
- Clustered transient MuSES nodal network



- Based on steady-state CFD
- Transient 30 minute run

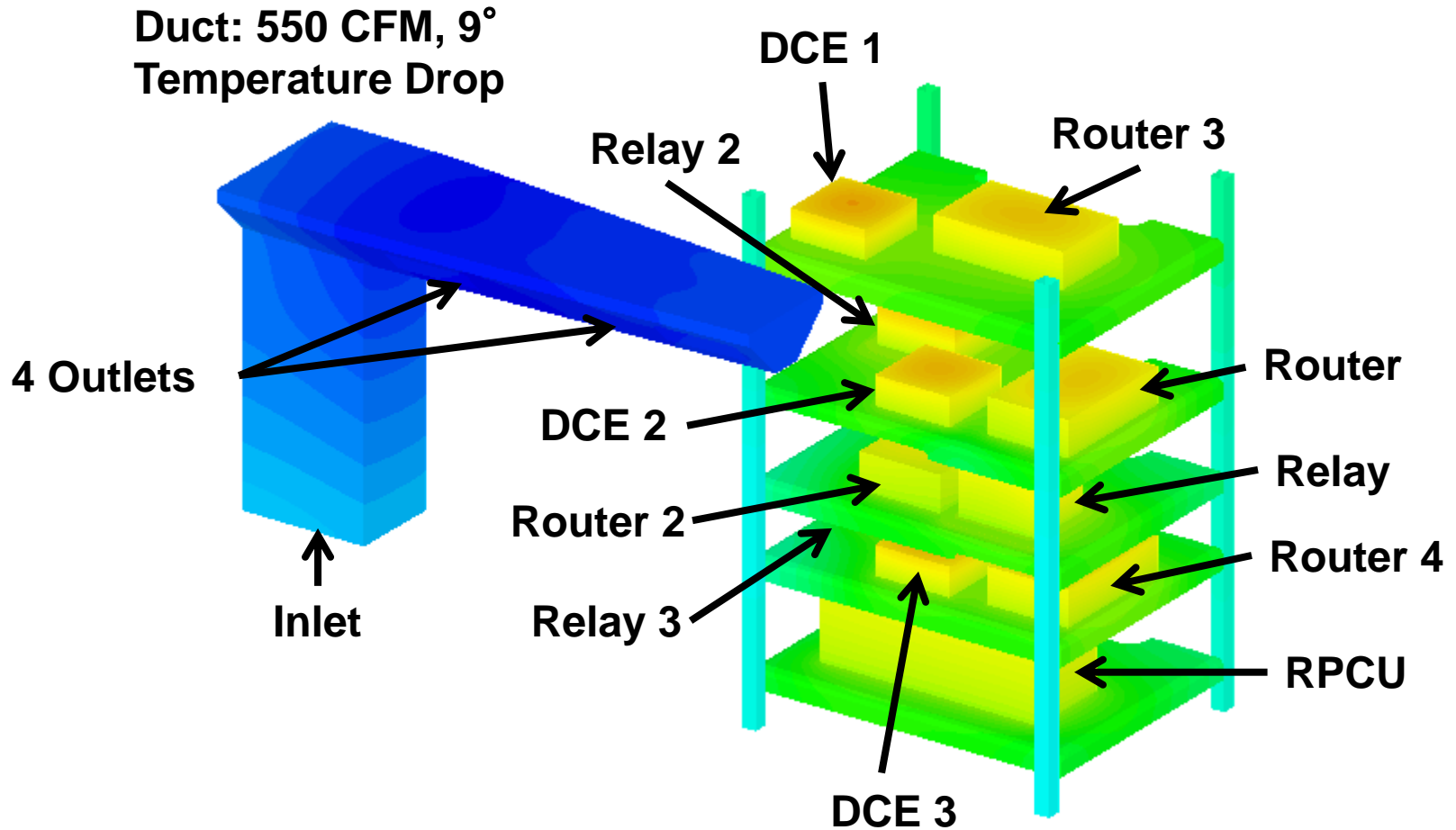
- Compare to fully transient CFD for 30 minutes
 - Extract average cell temperatures from clusters each minute

Validation Study



TECHNOLOGY DRIVEN. WARFIGHTER FOCUSED.

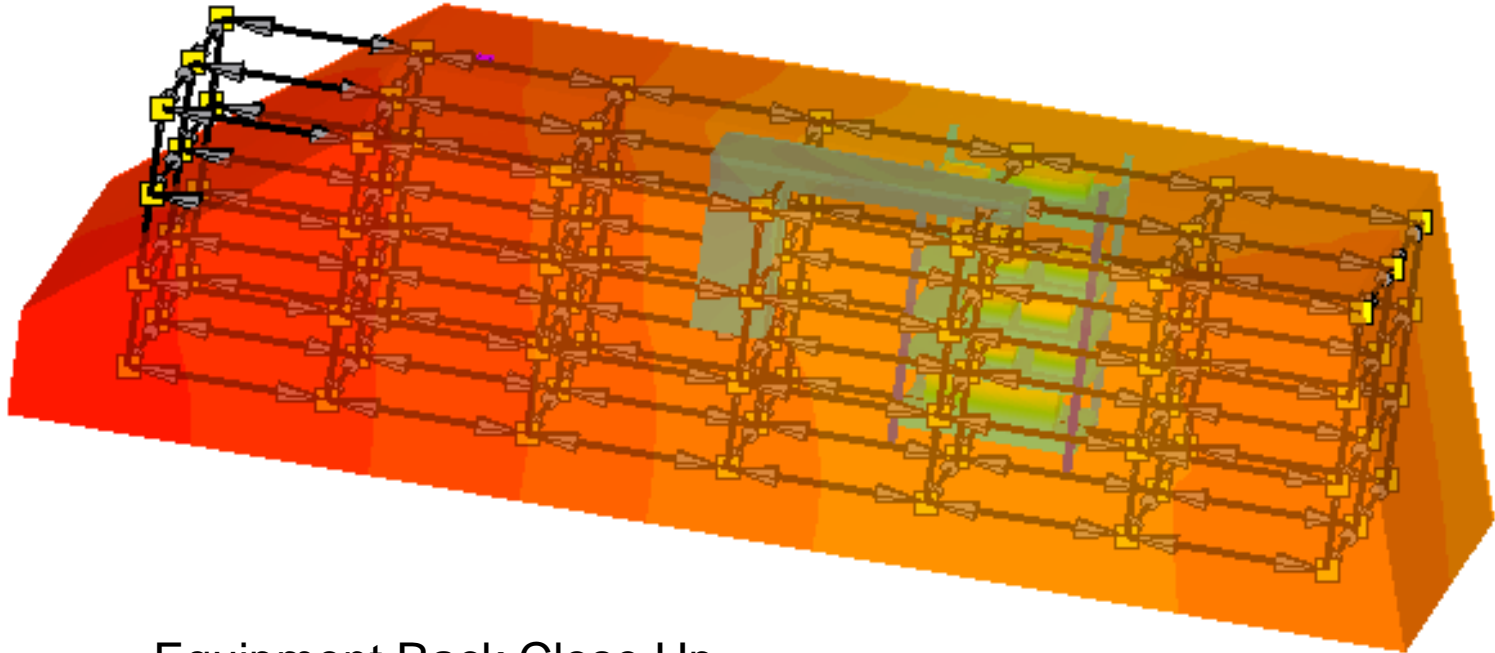
Equipment Rack Close-Up



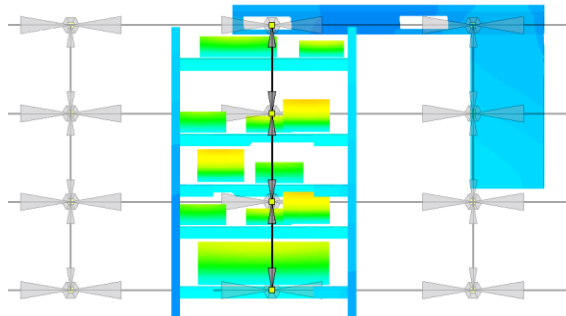
Subvolumed Nodal Arrangement



Subvolume nodal arrangement
72 Nodes Evenly Spaced



Equipment Rack Close Up

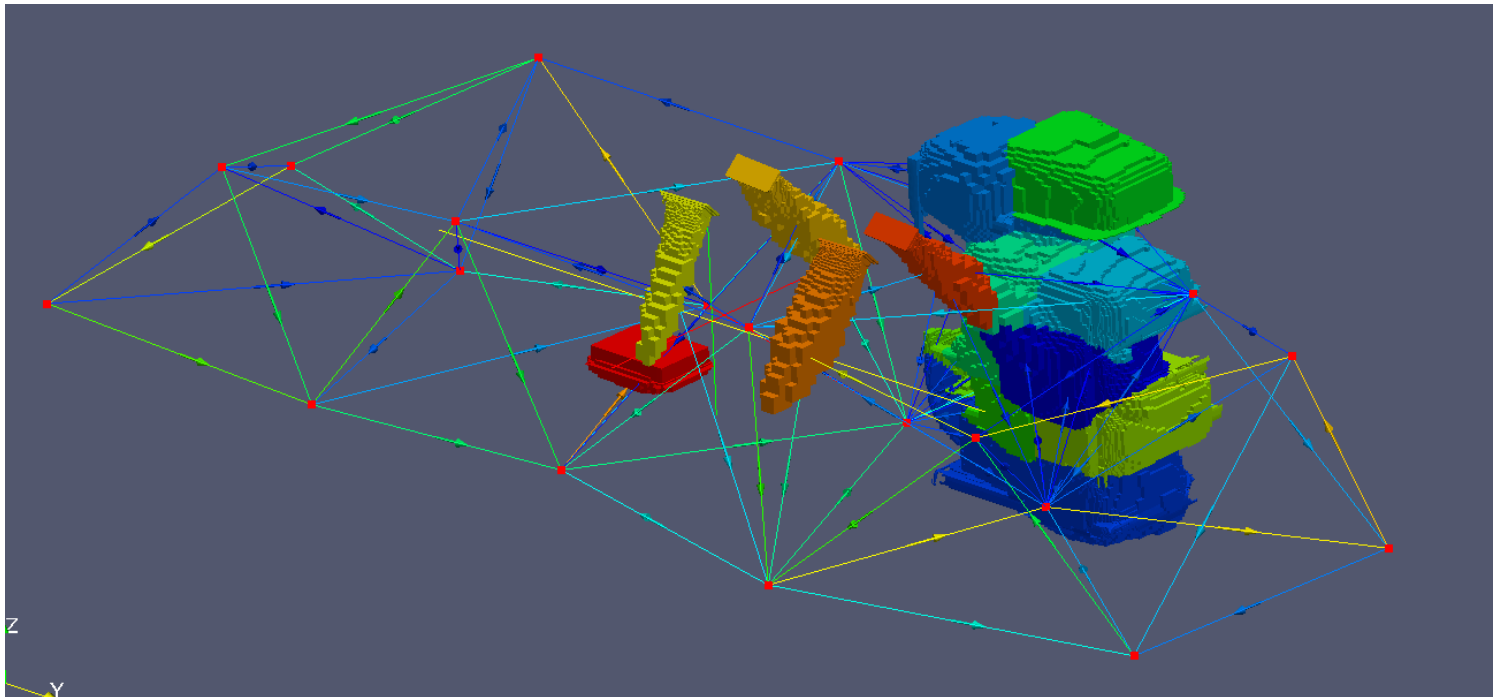


TECHNOLOGY DRIVEN. WARFIGHTER FOCUSED.

Clustering Nodal Arrangement

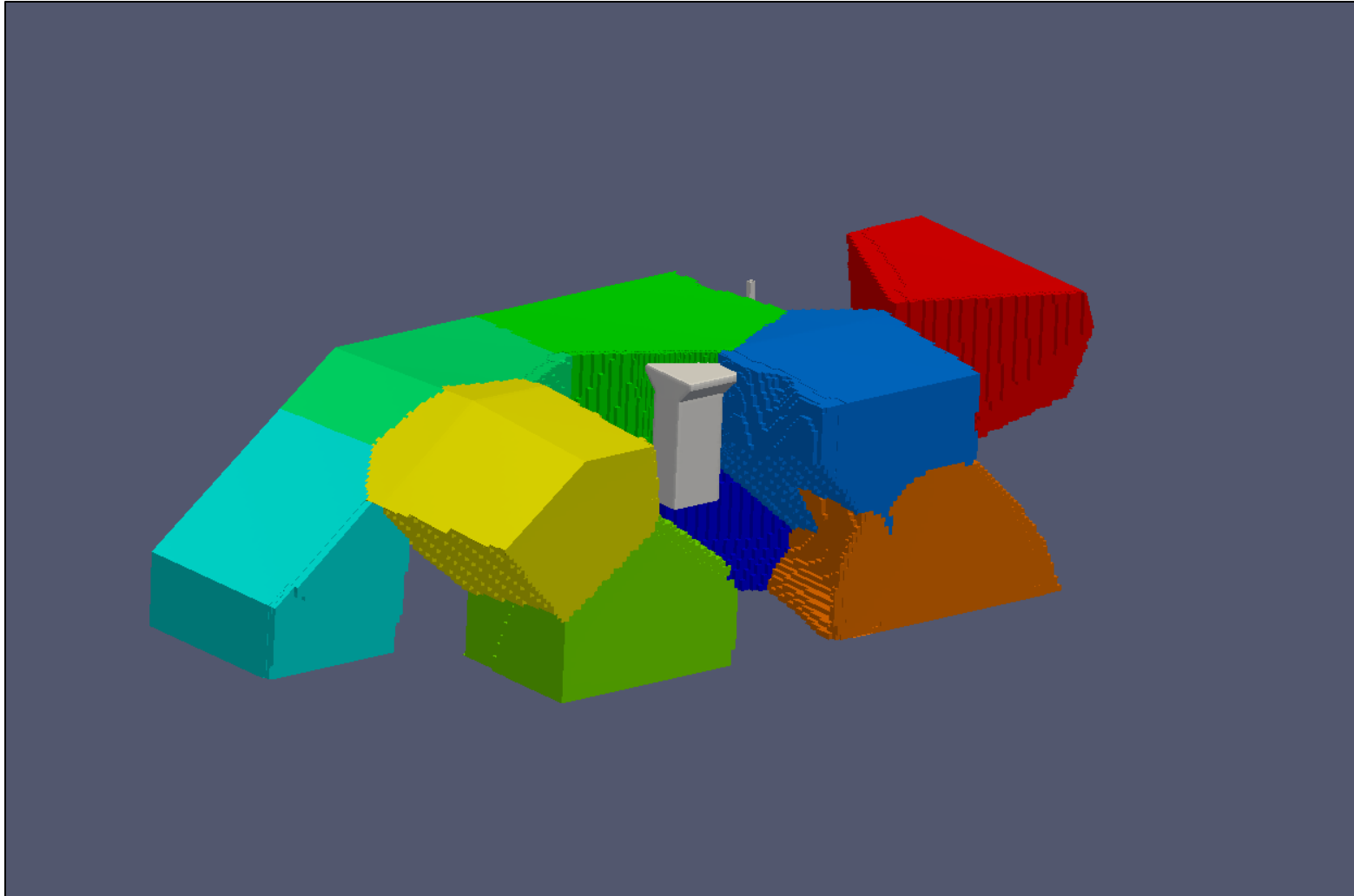


- 100% Euclidean
- Minimal time spent on clustering (not tailored)
- 35 Nodes
 - 10 around equipment
 - 5 around inlet/ outlets



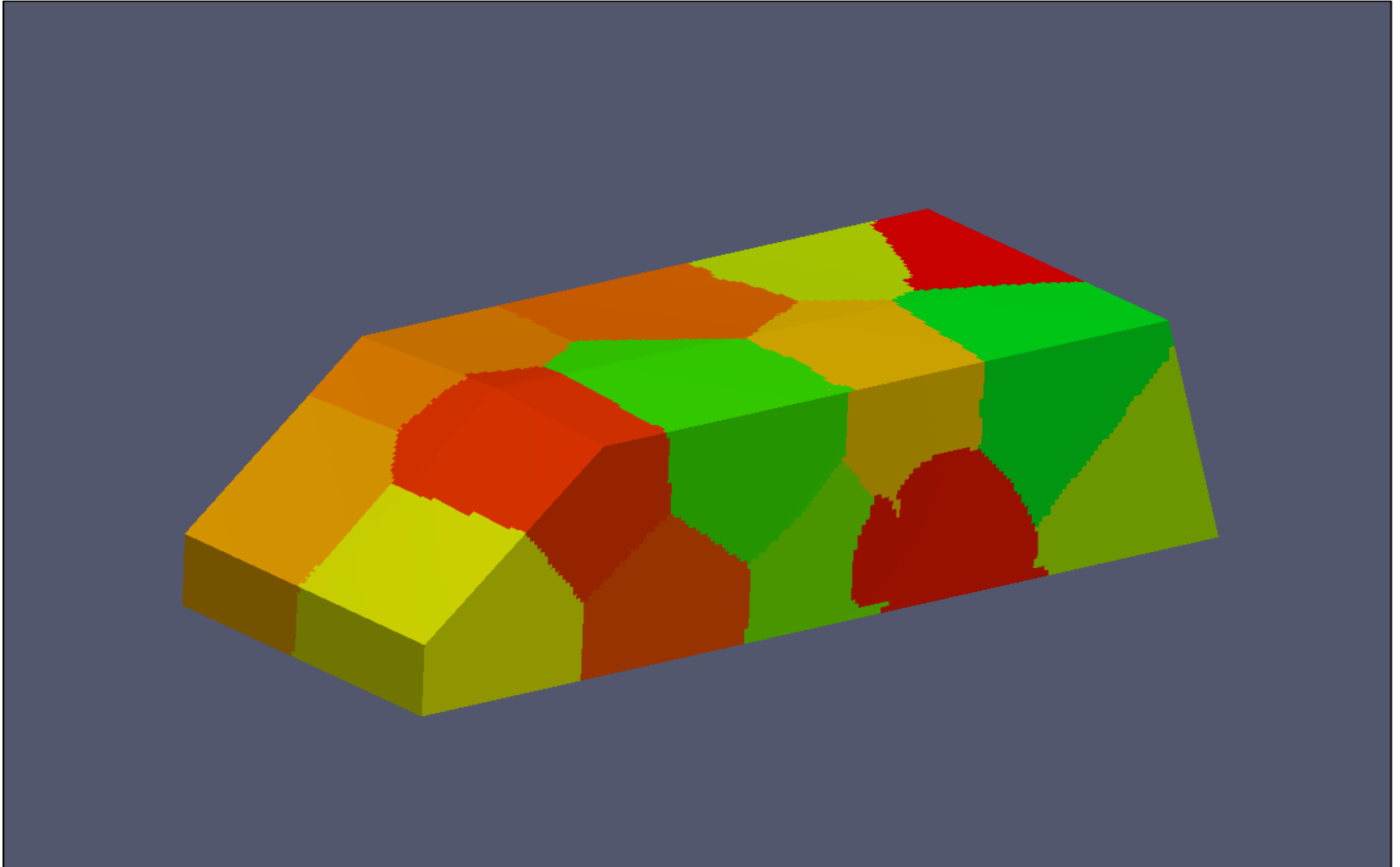
TECHNOLOGY DRIVEN. WARFIGHTER FOCUSED.

Validation Clustering View 2



TECHNOLOGY DRIVEN. WARFIGHTER FOCUSED.

Validation Clustering View 3



TECHNOLOGY DRIVEN. WARFIGHTER FOCUSED.

Clustered Network as Viewed in MuSES



File Edit View Tools Units Window Help

Model Browser

Geometry Editor Analyze Signature Post Process

Elapsed Time (min) 0
Time Step 0

Animate Speed: + - Repeat

Results Environment Plot/Export Notes BRDF

Part 2 Part Type: Calculated, 2 Sided

Average Temperature (°C)

Front	39.3094	<input checked="" type="radio"/> Display Heat Rates/BCs
Back	39.3272	<input type="radio"/> Display Heat Rates/BCs

Heat Rates Boundary Conditions

Display as Flux

Heat Rate (W)

	Incident	Outgoing	Net
Q Conduction			193.369
Q Convection	0	0	0
Q Radiation	12782.7 *	12877.2 *	-94.4512
Q Imposed			0

Model Size (mm):
X = 1648.14
Y = 4262.48
Z = 1068.98

00:00:00 Temperature

Visible Size (mm):
X = 1032.4
Y = 1128.5
Z = 943.551

Visible Counts:
Parts = 14
Elements = 49808

23.0 25.5 28.0 30.5 33.0 35.5 38.0 40.5 43.0

Display Thermal Results Temperature

Min Temperature (°C) 23 Display Scale Auto Scale Max Temperature (°C) 43

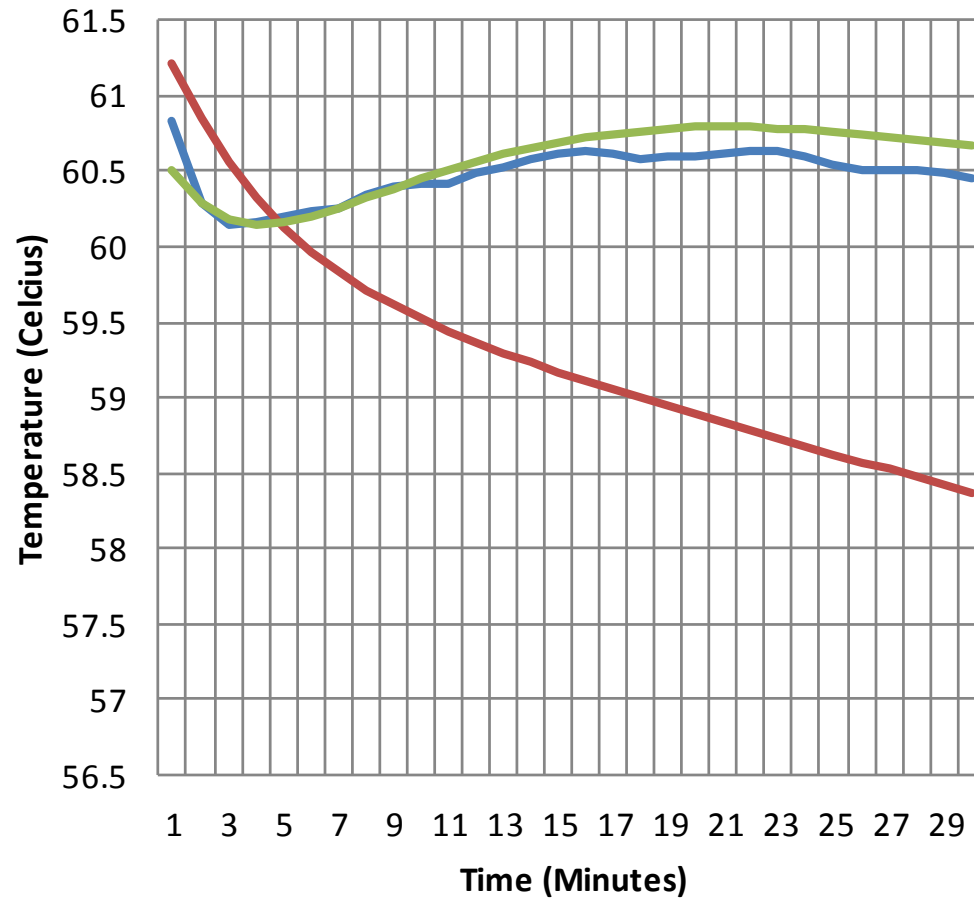
Messages

Irfs: 87
Quads: 49808

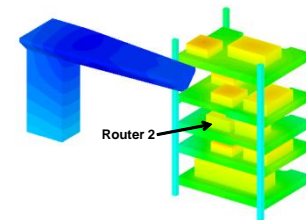
31) Thermal results data was written by version 10.2a-2011-14
32) Opened `/_home/rob/finaltest/killme_mined.tdf`
33) Using RGB graphics

TECHNOLOGY DRIVEN. WARFIGHTER FOCUSED.

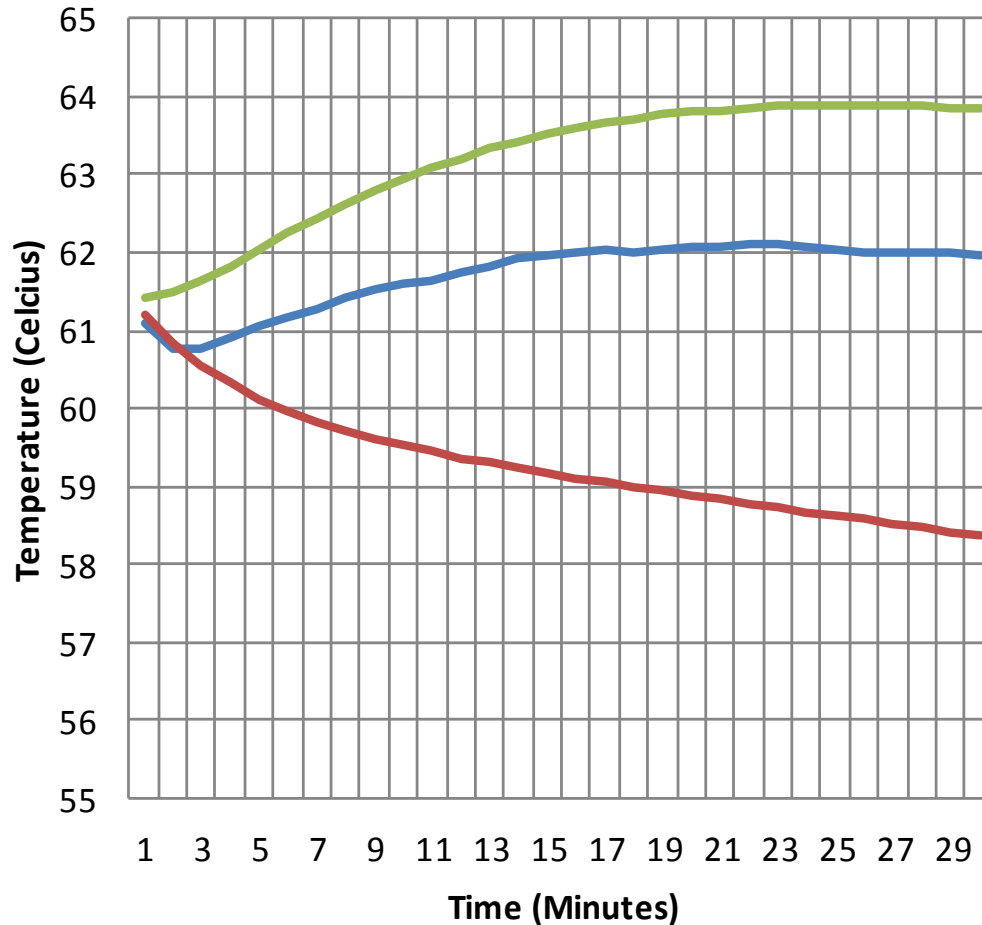
Router 2 Fluid



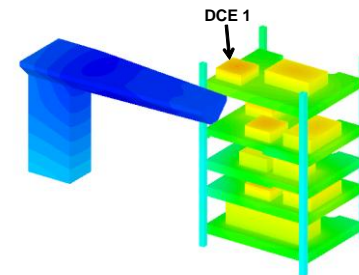
- CFD
- DCE1, Router3 Subvolumed
- Clustered

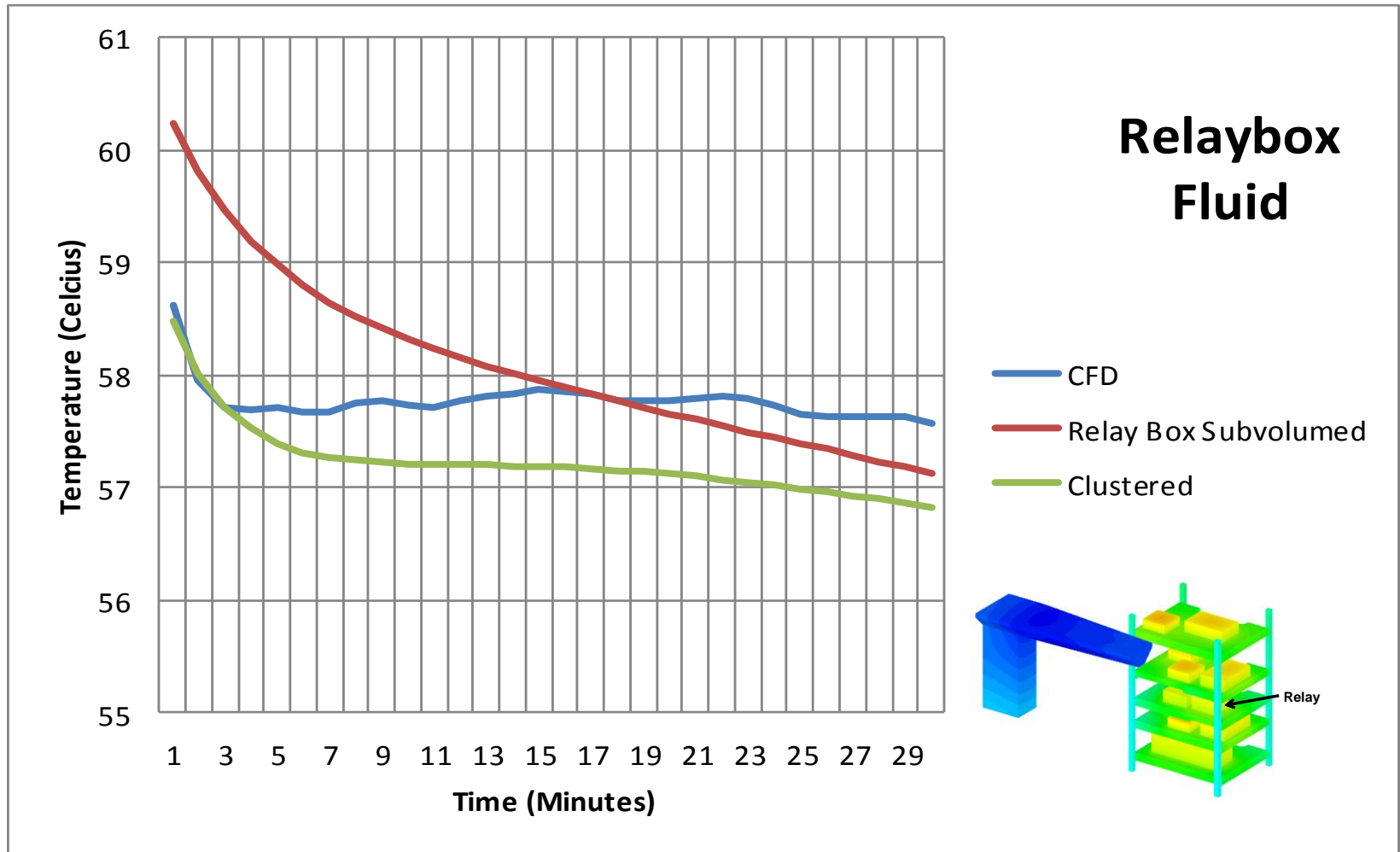


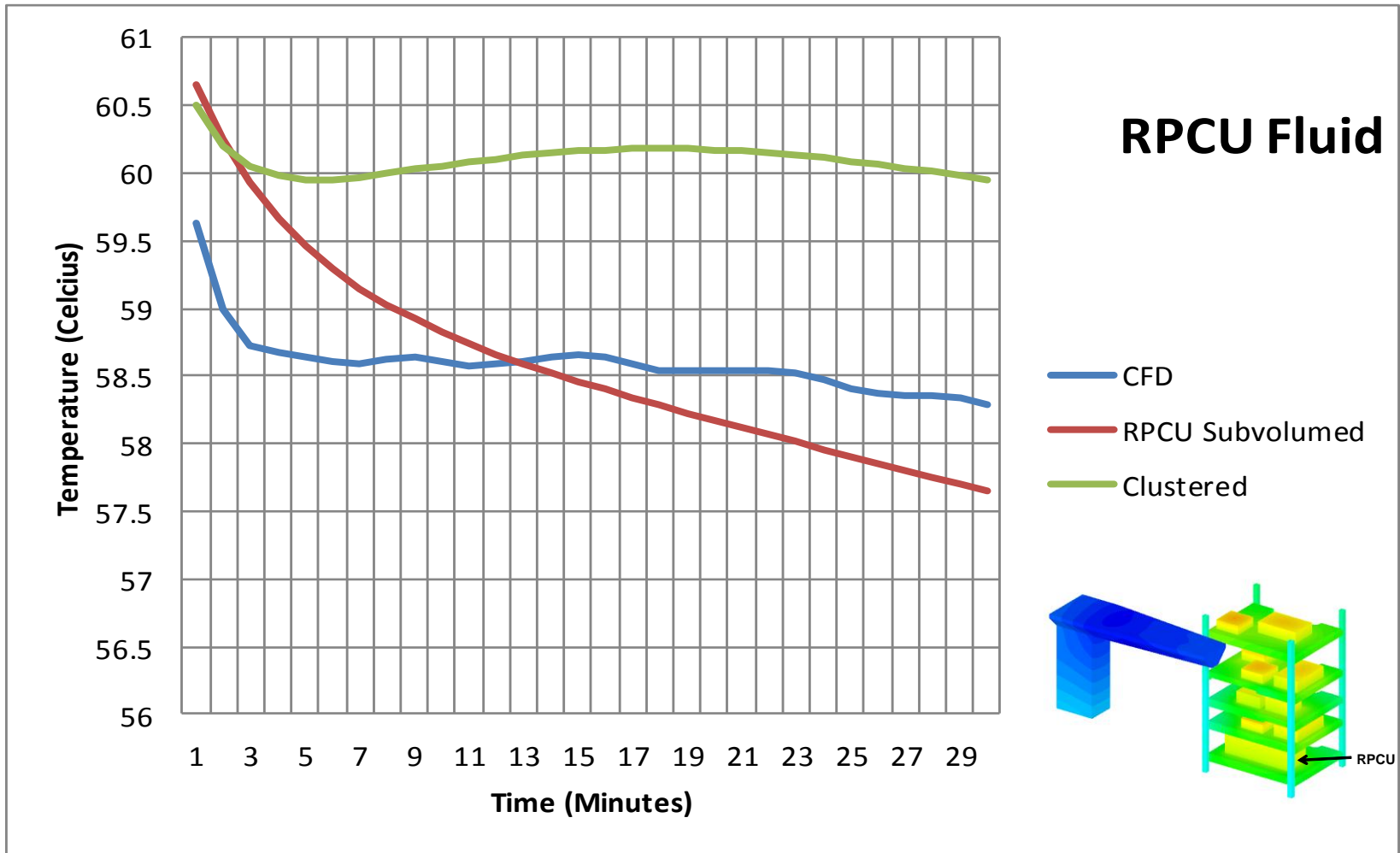
DCE1 Fluid

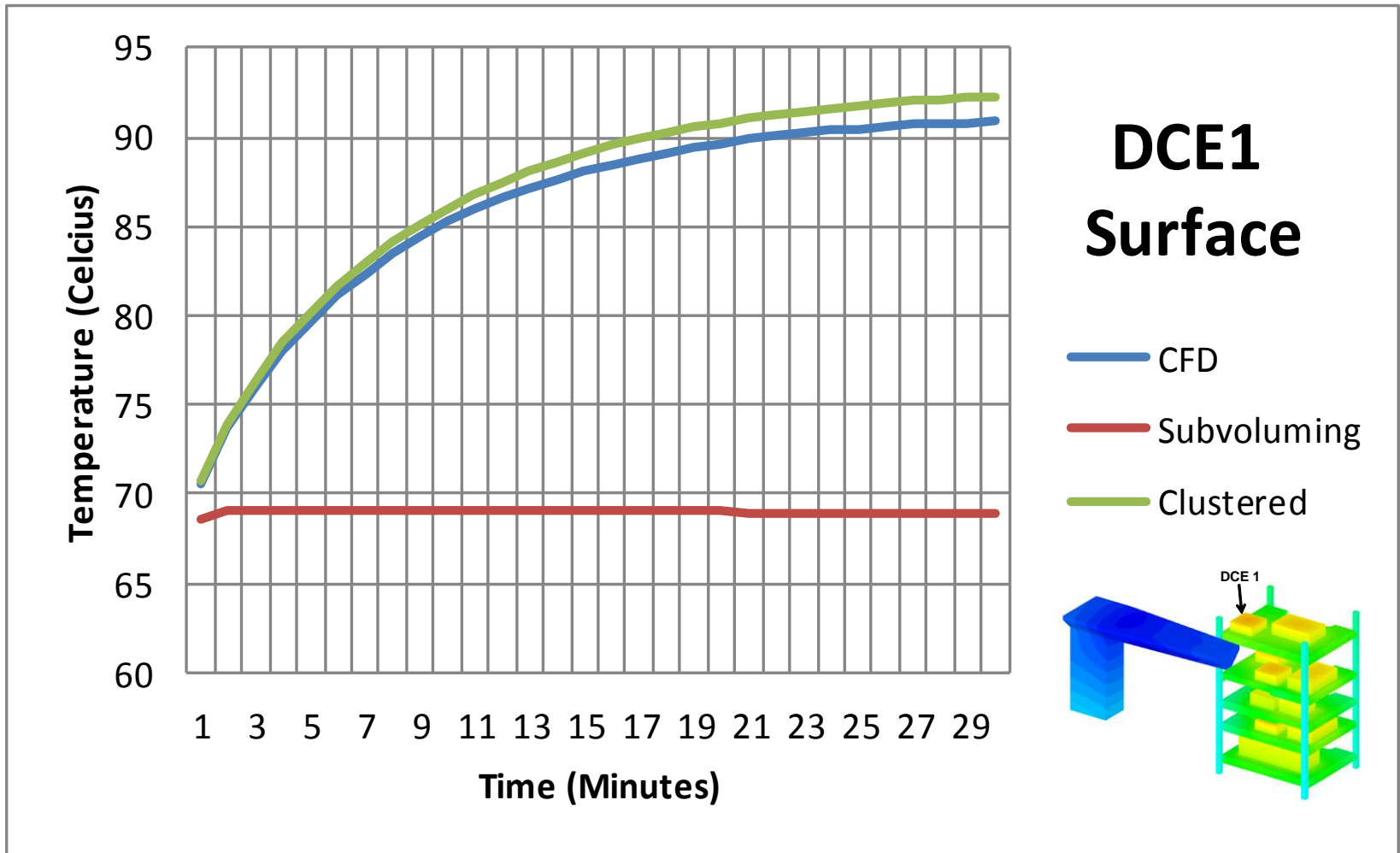


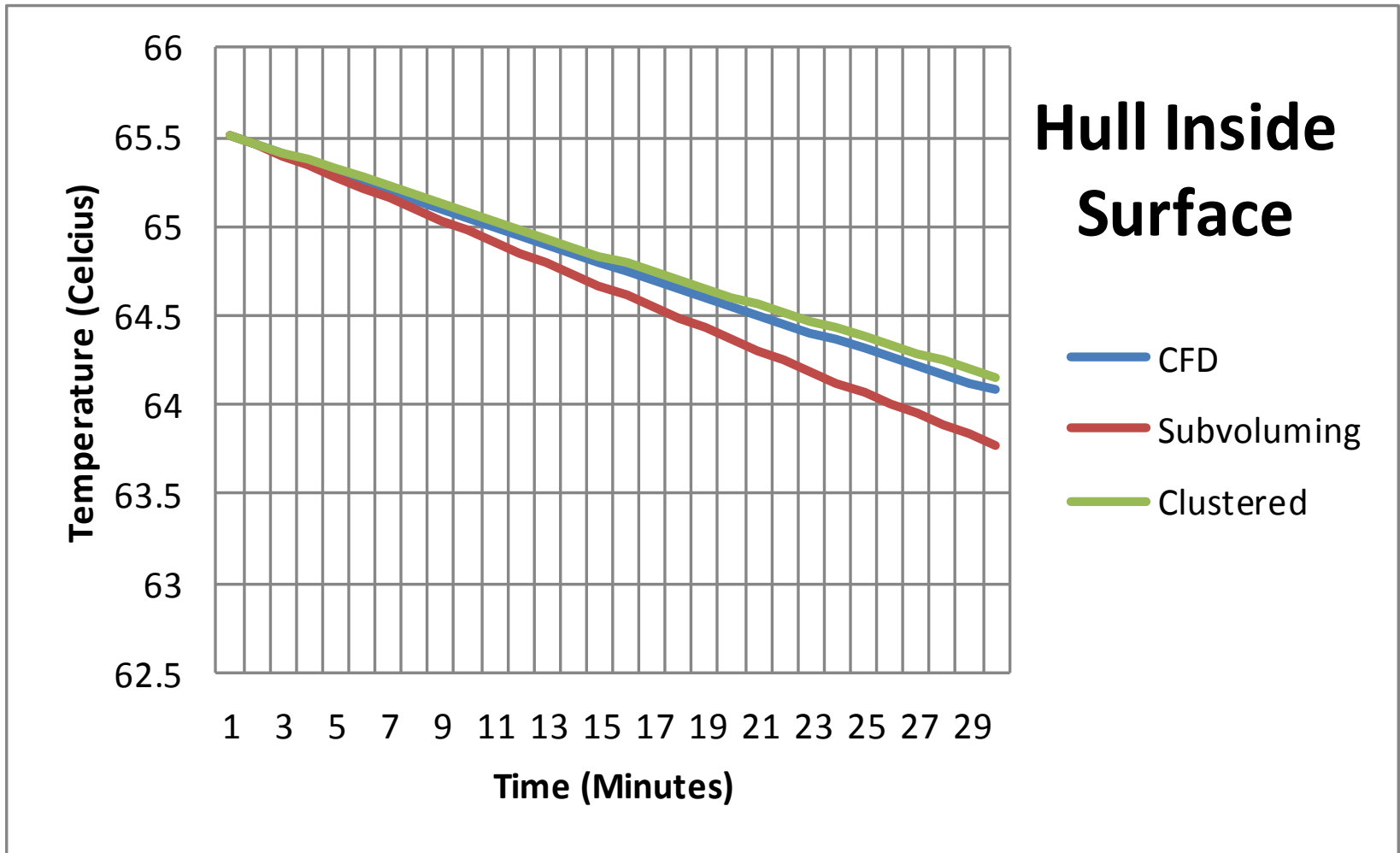
- CFD
- DCE1, Router3 Subvolumed
- Clustered







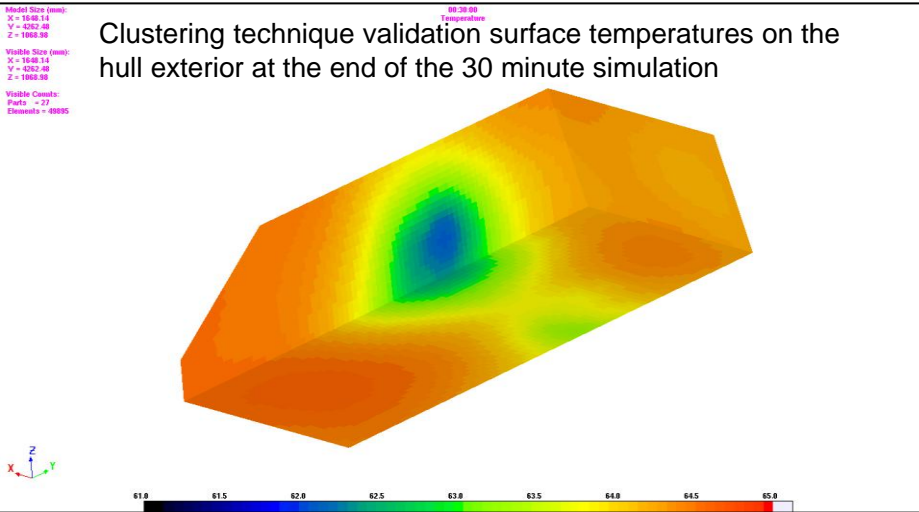




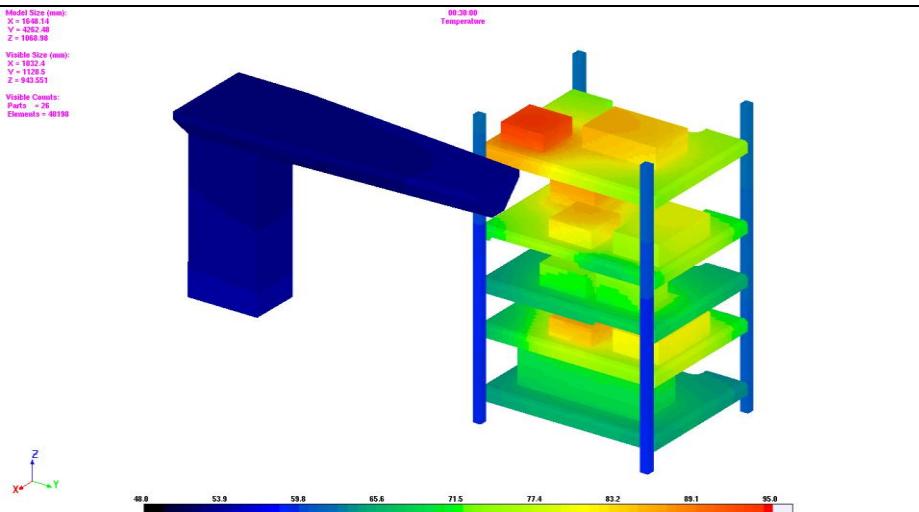
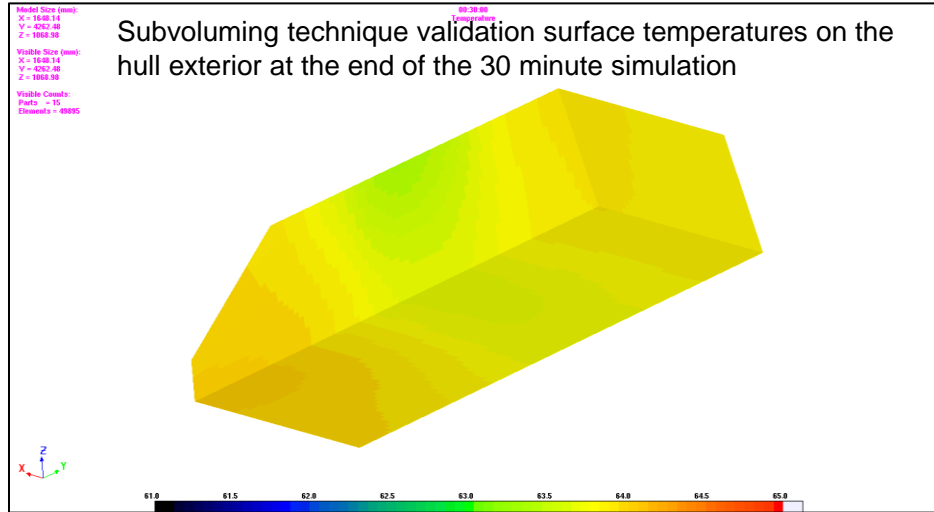
Demonstrated Need For Localized Convection



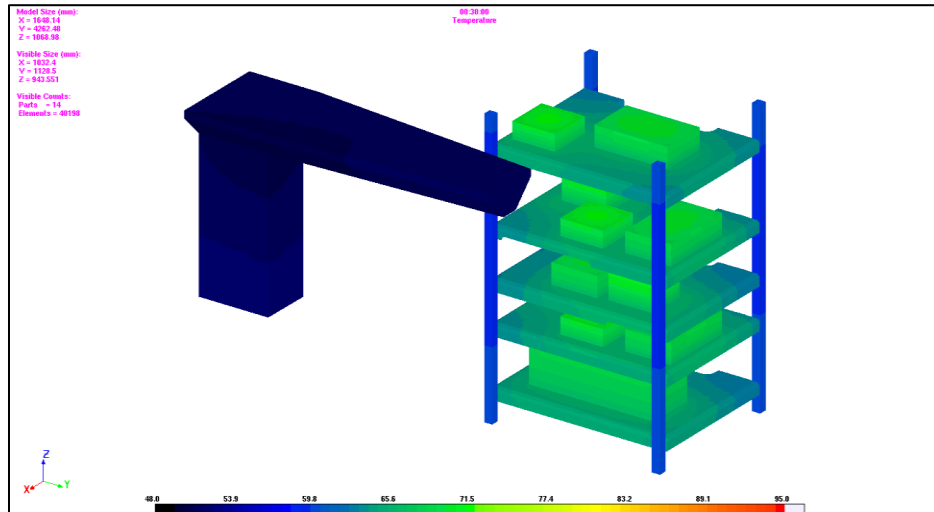
Clustering technique validation surface temperatures on the hull exterior at the end of the 30 minute simulation



Subvoluming technique validation surface temperatures on the hull exterior at the end of the 30 minute simulation



Clustering technique validation surface temperatures on the equipment rack at the end of the 30 minute simulation.



Subvoluming technique validation surface temperatures on the equipment rack at the end of the 30 minute simulation.

TECHNOLOGY DRIVEN. WARFIGHTER FOCUSED.

Remapping Problem



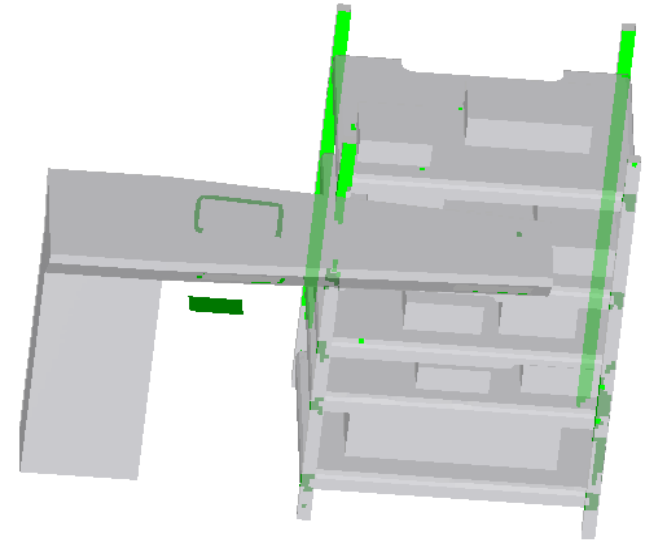
Sometimes:

$$h_{local} = \frac{h_{film}(T_{film}-T_{wall})}{T_{local}-T_{wall}}$$

produces a negative h_{local}

Options:

- Set to zero
- Use absolute value <-Recommended



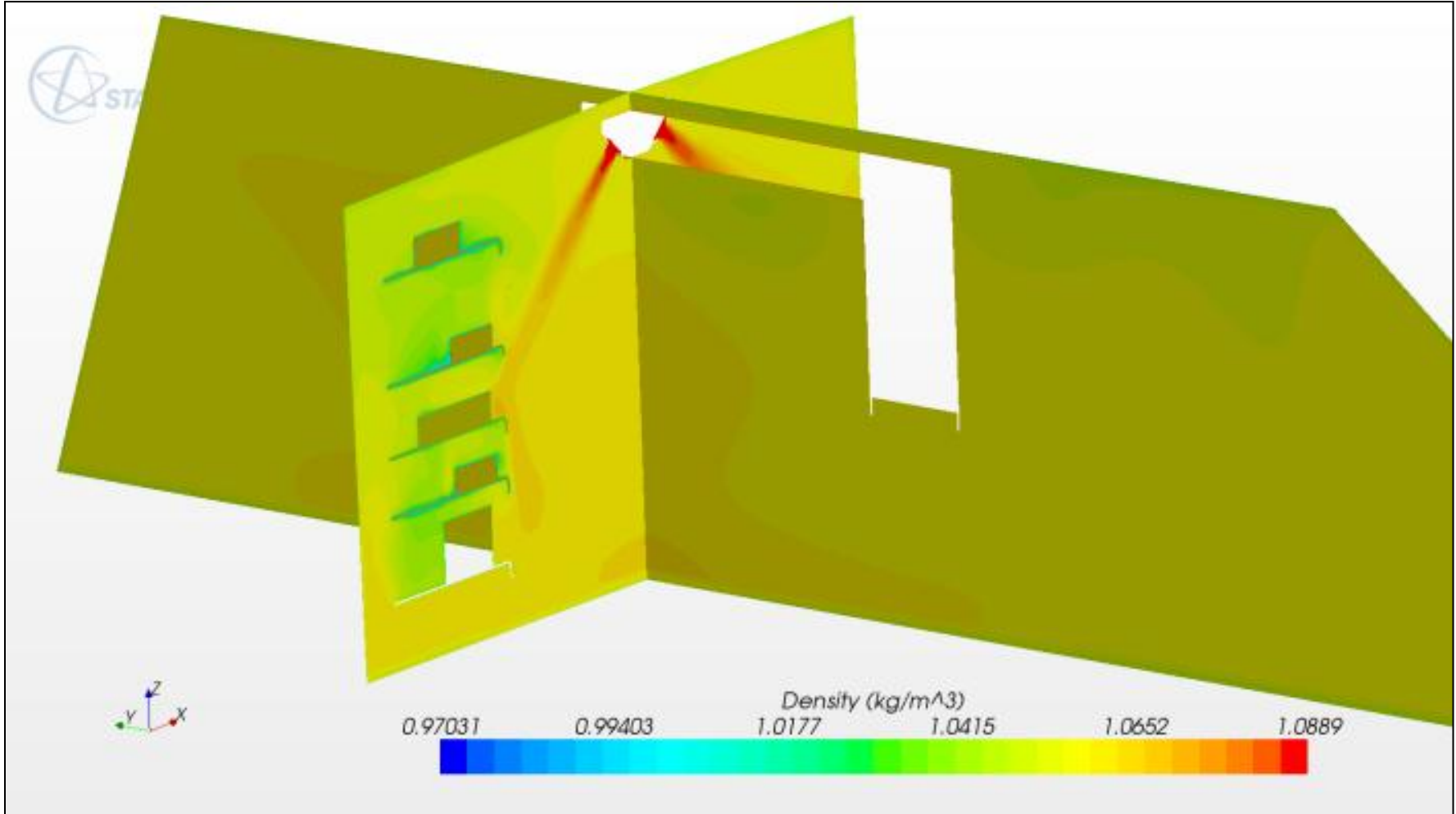
Validation Case Elements Which Have Negative Convection When Remapped

Conservation of Volume Problem



- It was noted the network is leaky for both subvoluming and cluster based networks.
 - Due to tolerances
 - Up to 10% imbalance seen on test case
 - Also should have been based on conservation of mass.
- Volume Balance Fix:
 - Start with the first cluster and assess all incoming and outgoing advection links. If there is a source outside the domain, enforce this in the calculation as well.
 - Divide the imbalance in half and apply half to the incoming and half to the outgoing flow to make them balance to zero (or reduce by a relaxation factor). Do this proportional to each advection link divided by the total incoming or outgoing flow.
 - Evaluate each cluster according to step 2 successively until each cluster exhibits no further imbalance.

Density Plot



TECHNOLOGY DRIVEN. WARFIGHTER FOCUSED.

Agenda



- About Rob E Smith
- Research Objectives
- Quick Introduction to Data Mining
- cfdMine Tool
- cfdMine Mahalanobis
- Validation Results
- Conclusions/ Next Steps

Conclusions



- The new overall process of clustering a steady-state CFD domain into a transient lumped fluid network simulation performs well
- It is absolutely critical to conserve mass within the fluid network
- Remapping CFD fluid film convection to localized fluid nodes is critical to an accurate solution
- Mahalanobis distance is better than Euclidean, but too computationally intensive for the present constraints
- Clustering is an effective method to track temperatures around specific equipment or locations
- Clustering provides a beneficial new way to visualize flowfields
- Subvoluming should be used with extreme caution

New Contributions to the Field



- First to implement clustering to simplify a CFD network into a simplified network
- First known volume weighted clustering of a CFD domain.
- First known remapping of convection coefficients were remapped to localized nodes in a network
- First use of fixed nodes where the nodes were fixed at areas of interest during clustering for the purpose of a simplified nodal network.
- Proposed new methods to stabilize the Mahalanobis distance and make it usable.
- Showed several new ways to visualize flow data

Future Work



- Implement conservation of mass
- Investigate use of unidirectional flux linkages
- Look into enhanced subvoluming
 - Control volumes around equipment
 - Localized convection
- Investigate importance of transient versus steady-state convection
- Convection “view factors”
 - Solve a steady-state CFD solution
 - Release particles from all wall elements and boundary elements (perform a particle trace)
 - Account for the mixing and the fraction of volume of influence each wall-released particle
 - Ultimately track the influence each wall element / boundary has on each other element/ boundary

LOONEY TUNES

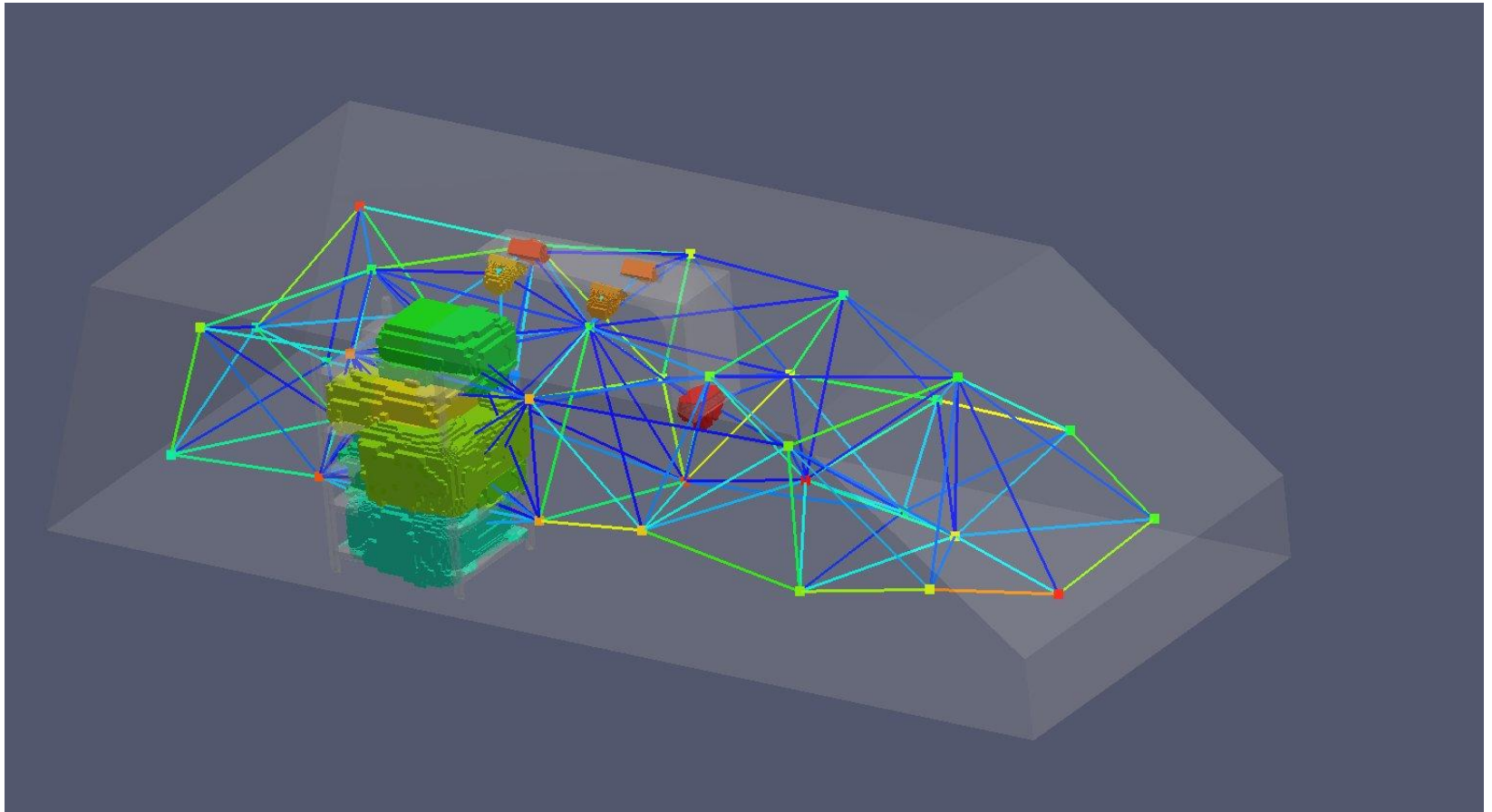


QUESTIONS?



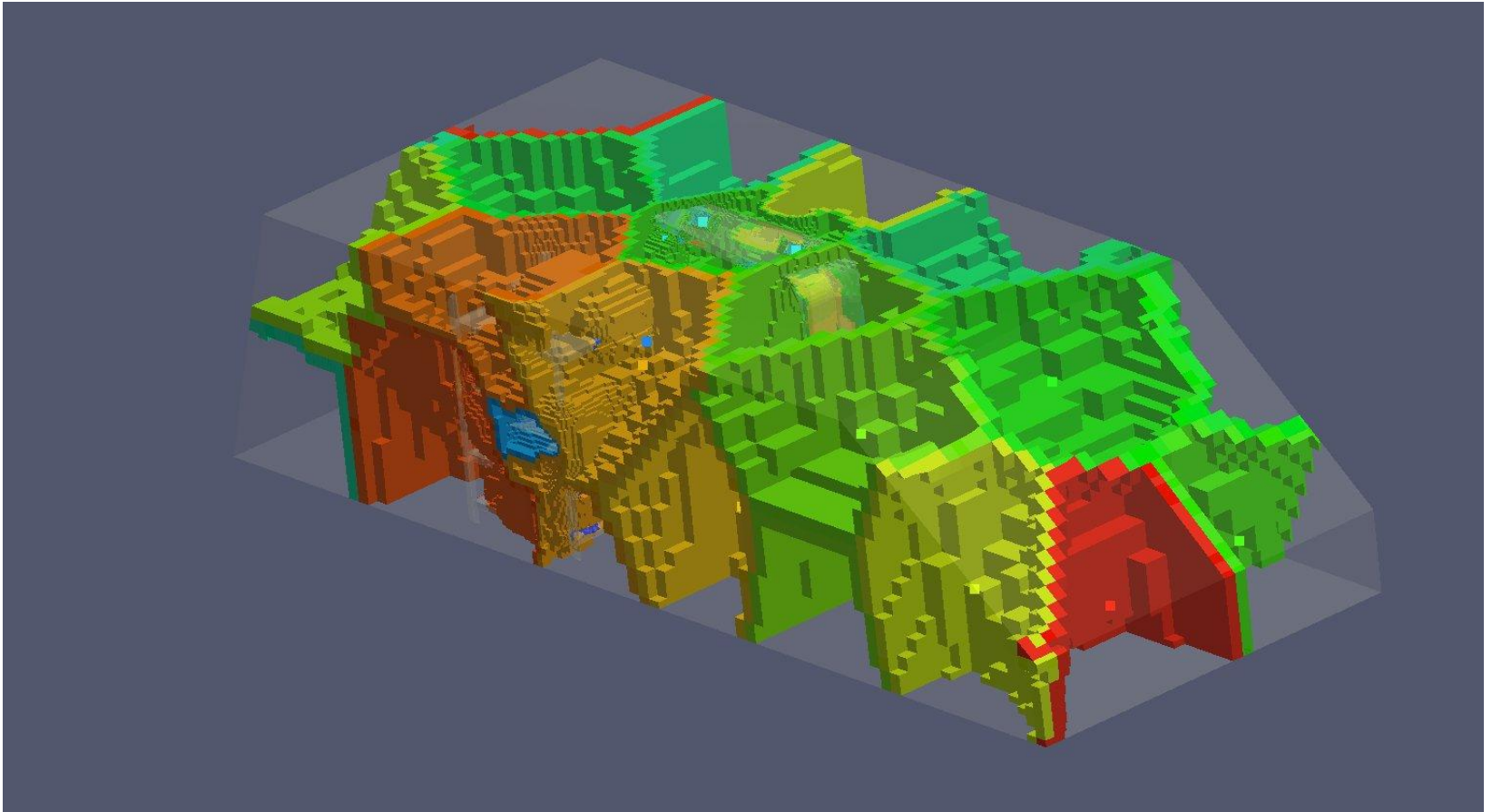
BACKUP SLIDES

50 Nodes



TECHNOLOGY DRIVEN. WARFIGHTER FOCUSED.

50 Nodes



TECHNOLOGY DRIVEN. WARFIGHTER FOCUSED.

K-Means

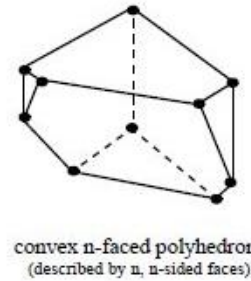
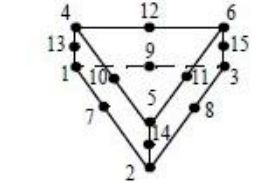
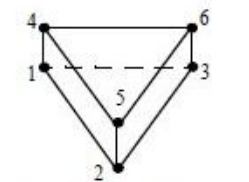
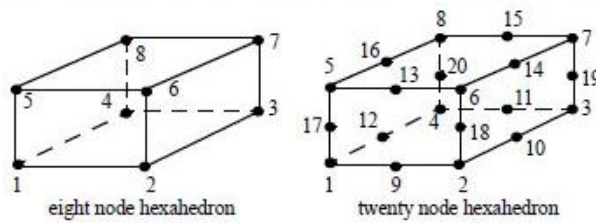
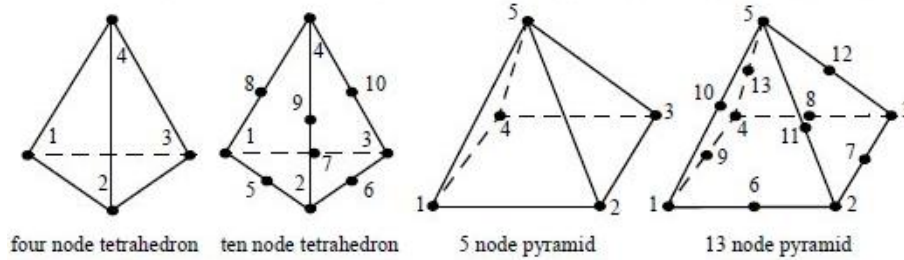
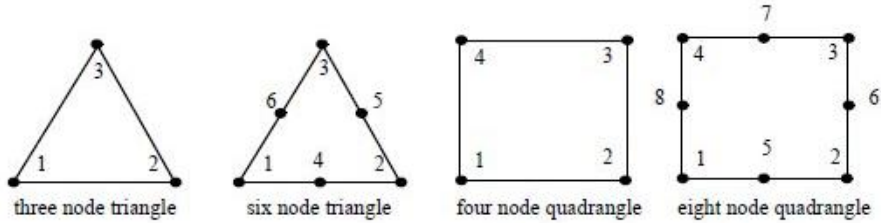
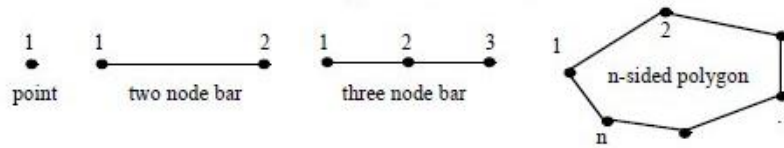


- K-Means was first suggested in 1962 by Sebestyen [15] and independently in 1967 by MacQueen [16]. A detailed history and full pedantic discussion of various deviations may be found in "K-Means Clustering: A Half-Century Synthesis" [17].
- Performance function: Minimize
 - *Sum Squared Error* = $\sum_{l=1}^K \sum_{x \in S_l} \|\vec{x} - m_l\|^2$
- K-Means Iterative Function
 - $\bar{m}_l = \frac{1}{n_l} \sum_{x \in S_l} \vec{x}$

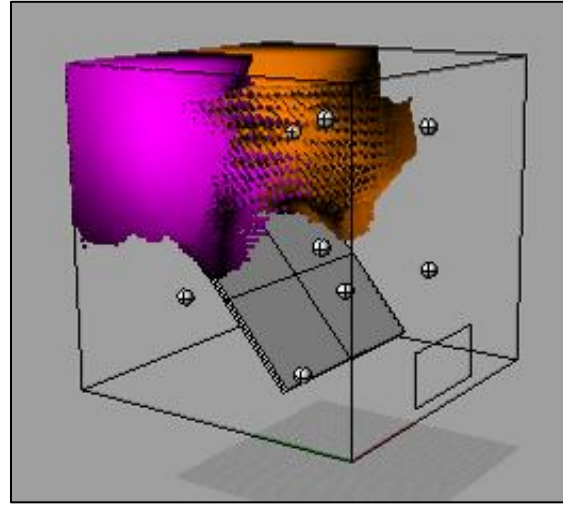
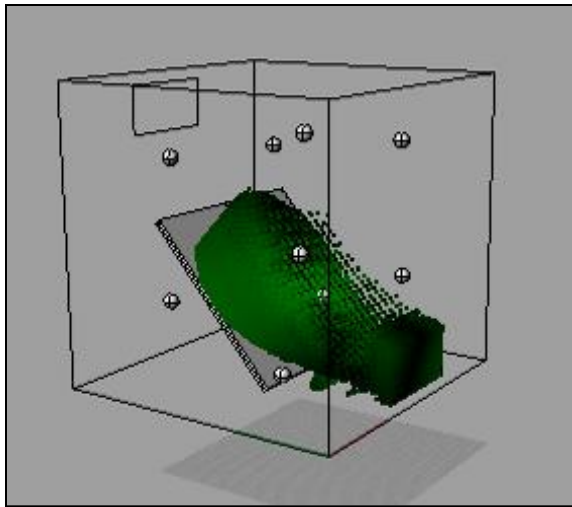
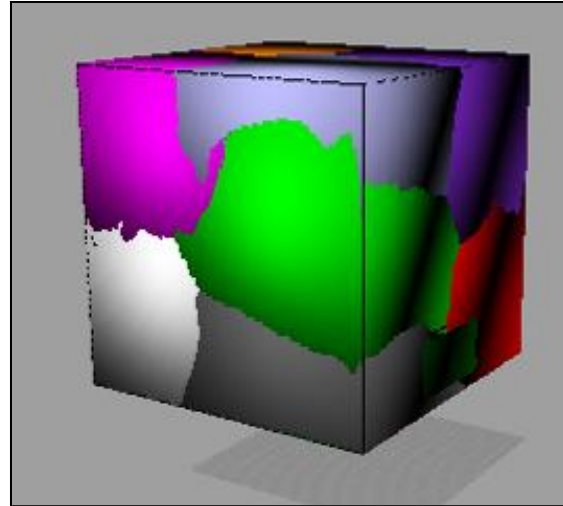
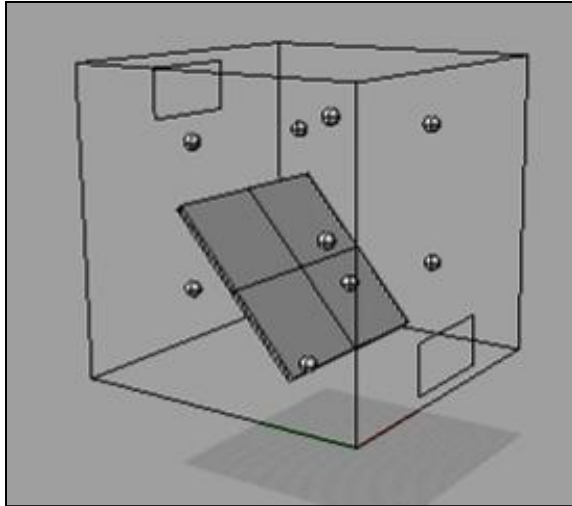
K-Means Basic Algorithm



- for $i=1$ to number of tries
 - repeat
 - randomly select cluster nodes
 - compute distance of every point in domain to nodes
 - assign points to cluster based on minimum distance
 - compute volume weighted means to find new nodes
 - track the solution with minimum total error
 - until no change
- end for

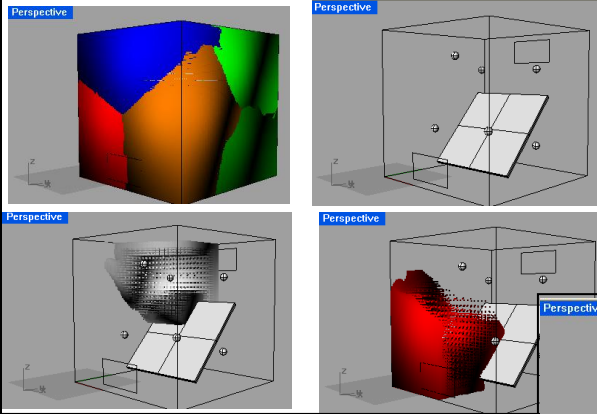


Example Clustering

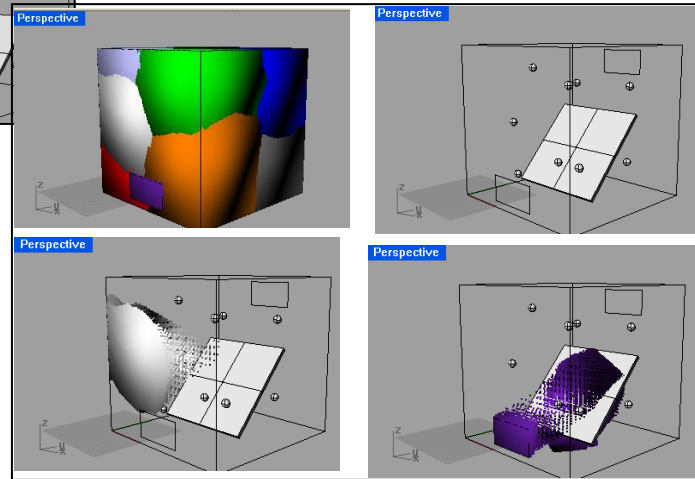


TECHNOLOGY DRIVEN. WARFIGHTER FOCUSED.

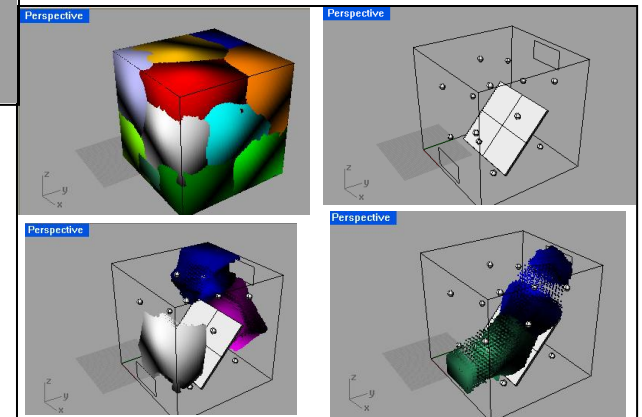
Number of Clusters



7 Clusters



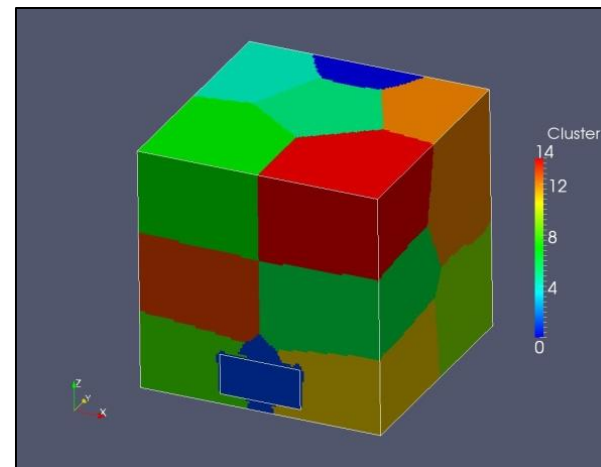
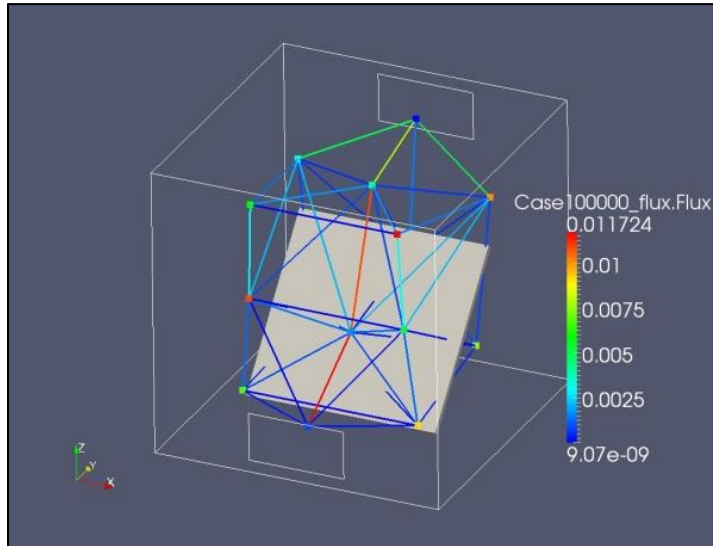
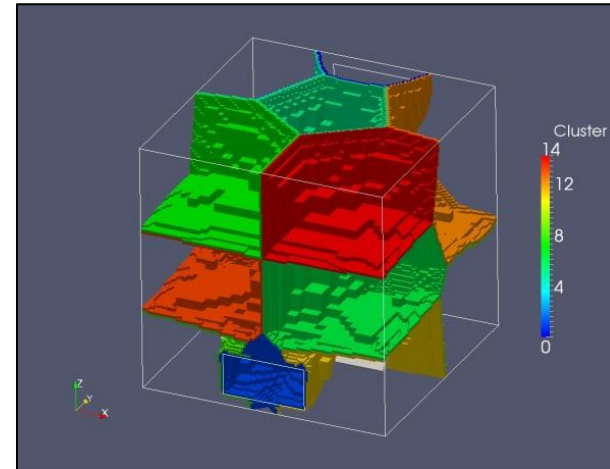
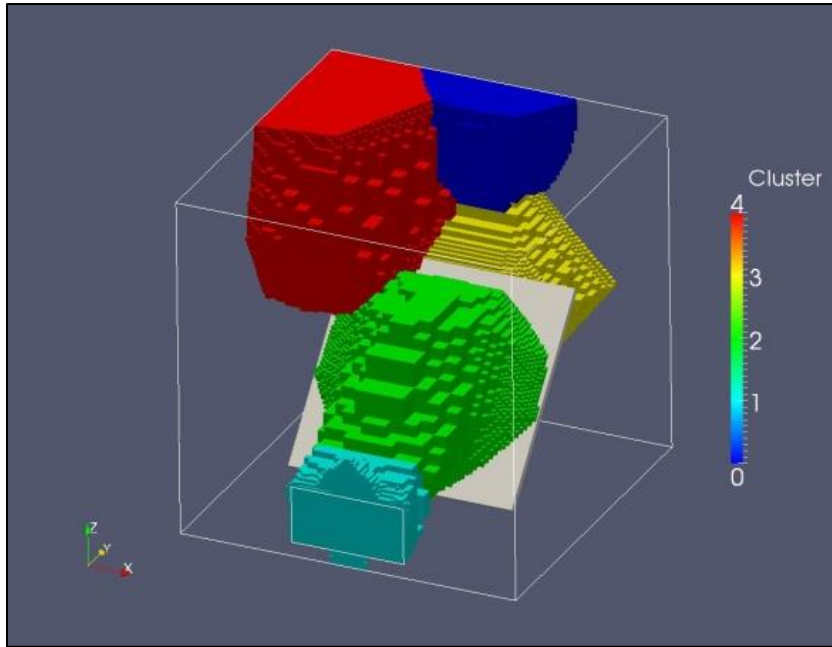
10 Clusters



15 Clusters

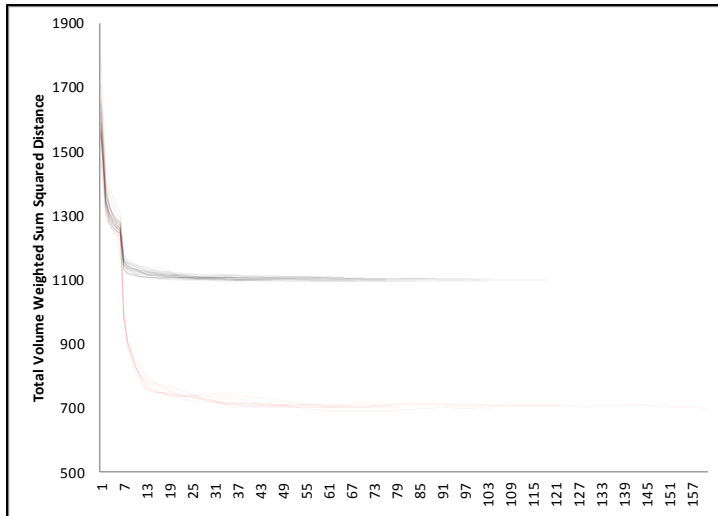
Result: 10 clusters is the minimum for this particular case to capture steam.

Visualization of Test Case 1

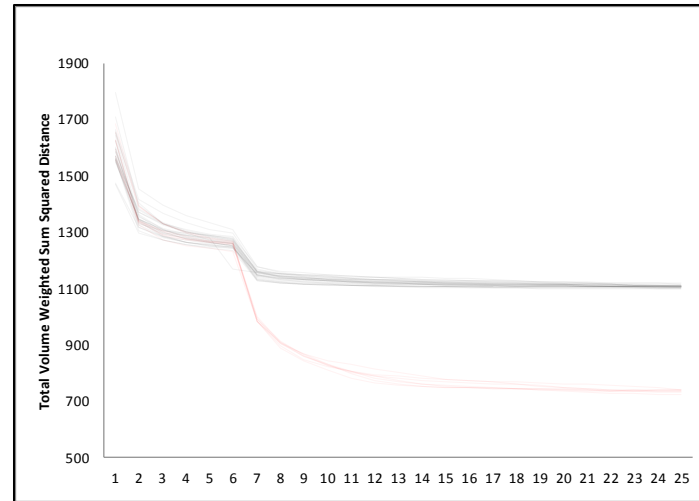


TECHNOLOGY DRIVEN. WARFIGHTER FOCUSED.

K-Means Mahalanobis Convergence



Typical k-means convergence plot for 33 random start trials on a vehicle duct geometry. The lines are translucent. The black lines show 25 runs with a 50% Mahalanobis/Euclidean blend versus the 8 red lines which show 75% Mahalanobis. Note the first 5 iterations are pure Euclidean



“Zoomed-in” plot of typical k-means convergence plot for 33 random start trials on an vehicle duct geometry. The lines are translucent. The black lines show 25 runs with a 50% Mahalanobis/Euclidean blend versus the 8 red lines which show 75% Mahalanobis. Note the first 5 iterations are pure Euclidean

TECHNOLOGY DRIVEN. WARFIGHTER FOCUSED.

Data Mining / Clustering



- Primary two categories for data clustering are:
 1. Hierarchical
 2. Partitional.
- Hierarchical clustering is based on the notion that objects are more related to nearby objects than to objects farther away.
 - Hierarchical clustering can be computed bottom-up or top-down. Hierarchical agglomerative clustering is starting with single elements and aggregating them into clusters.
 - Hierarchical divisive clustering is starting with the complete data set and dividing it into partitions.
- Partitional clustering operates on data globally
 - Exemplified by k-means where given an initial set of “k” means, points are assigned to the most similar “k” cluster.
 - Centroid-based clustering optimization is ultimately considered an np-hard problem, but is possible to solve for most practical cases.

New cfdMine Mahalanobis Algorithm



Mahalanobis cannot be used directly with K-means. Adds 36 unknowns and drops into non-optimal local minima even faster.

1. Pick random cluster centers from the domain and assign elements randomly
2. Z-normalize all the data in the domain
3. First m iterations are Euclidean
4. If iterations $> n$, compute the covariance matrix, otherwise identity matrix
5. Divide the x,y,z 3×3 inverse covariance matrix by the maximum value and multiply by the weight of importance for physical distance. For temperature, velocity magnitude, and velocity direction, divide the associated rows and columns by the diagonal value.
6. Blend the inverse covariance matrix with an identity matrix with variable weights on the diagonal.
7. Compute the Mahalanobis Distance from each domain point to the covariance matrix.
8. Assign each point the closest cluster.
9. Go to step 3 until converged

TECHNOLOGY DRIVEN. WARFIGHTER FOCUSED.