Bloomenergy

Moving Away from the Grid - Implementing Fuel Cell Technology for Data Centers

North American Network Operator Group

San Francisco
June 1st, 2105

Peter Gross, PE
VP, Mission Critical Systems
Bloom Energy

How It Works

Convert fuel directly into electricity without combustion

Conventional Electrical Generator

Natural Gas
(Chemical Energy)

↓

Combustion (Thermal Energy)

↓

Turbine
(Mechanical Energy)

↓

Alternator
(Electrical Energy)

Bloom's Fuel Cell

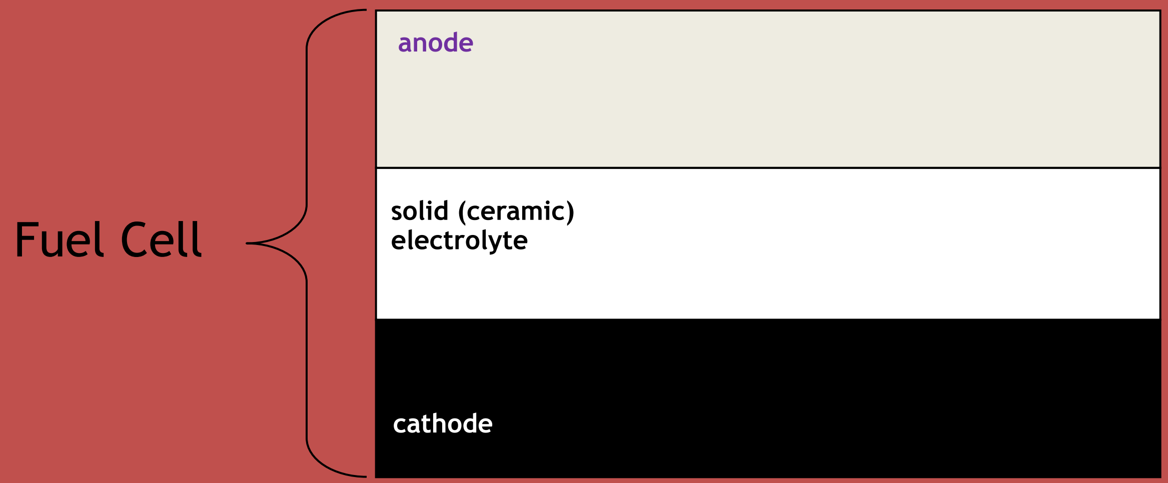
Natural Gas
(Chemical Energy)

↓

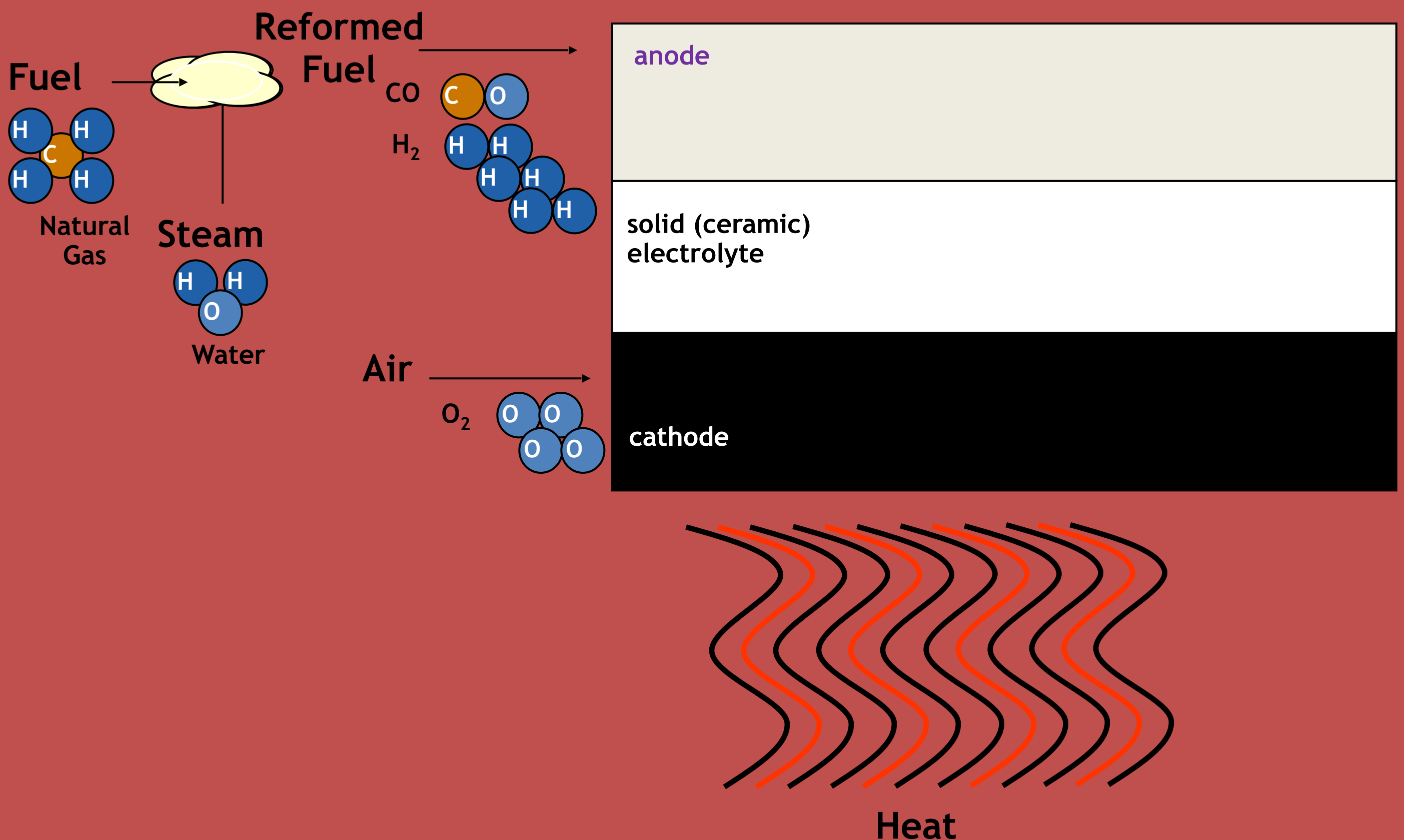
Electrochemical Reaction
(Electrical Energy)

Efficient = More Electricity

How a Fuel Cell Works

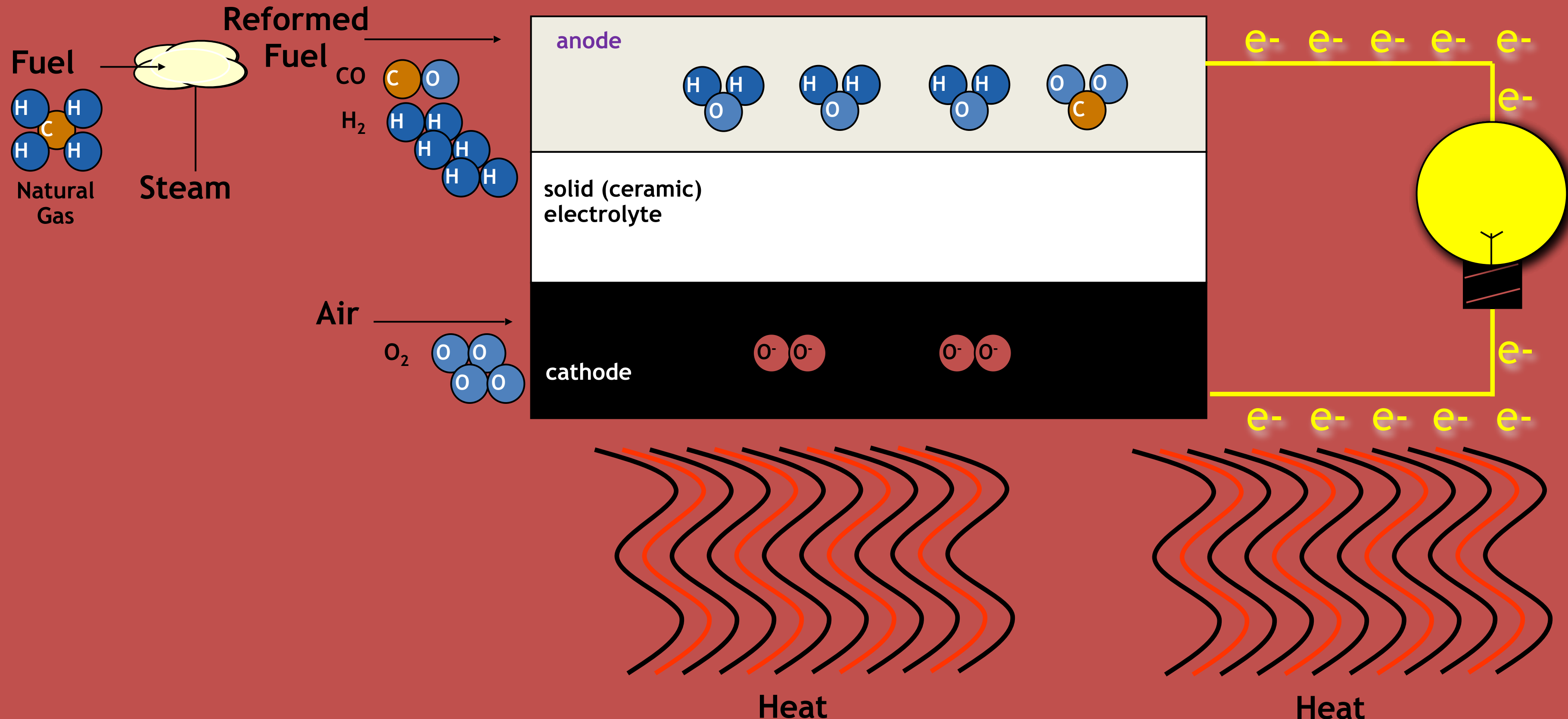


An electrochemical reaction converts
fuel and air into electricity… without combustion
metals).

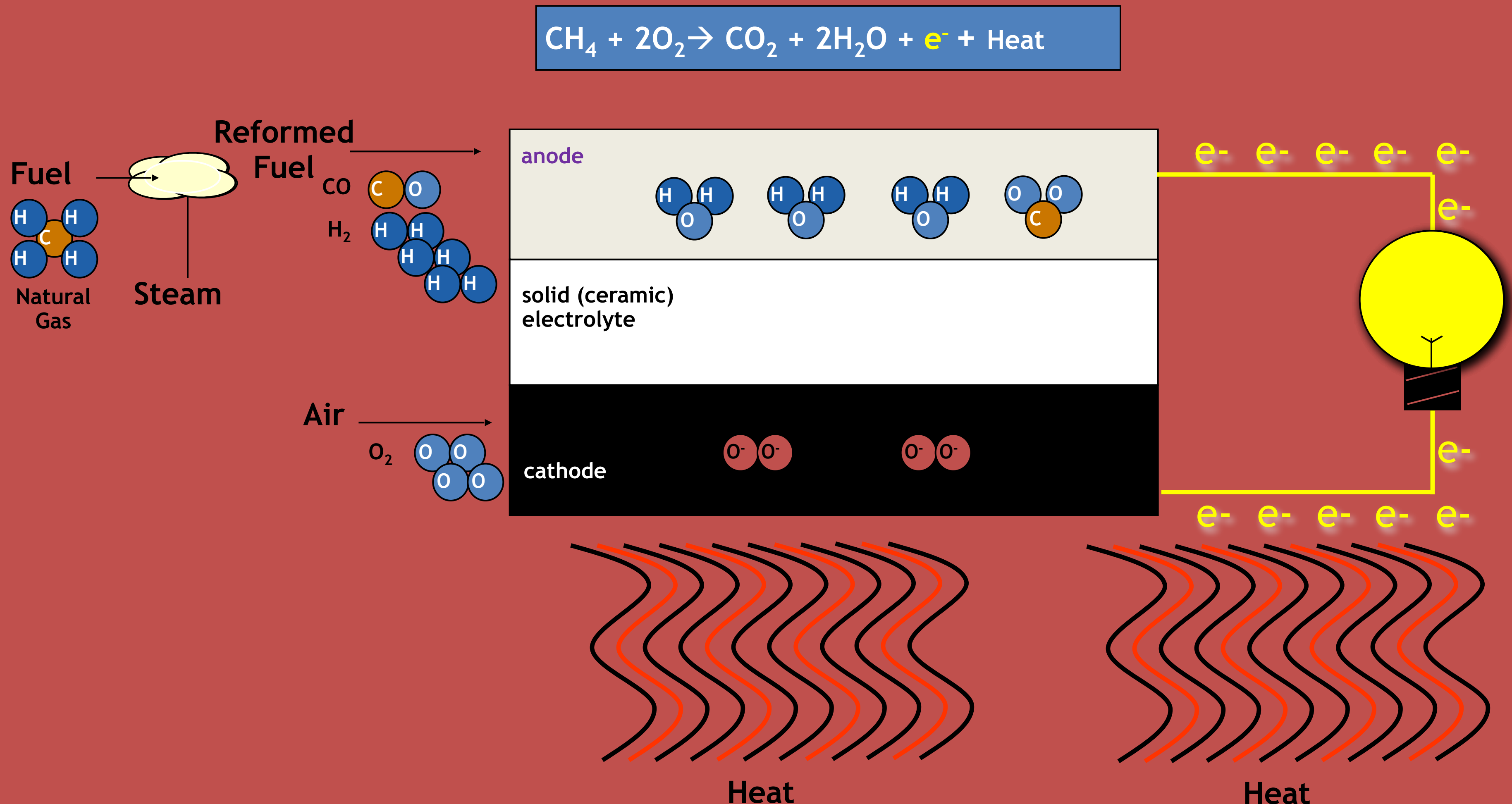


And steam mixes with fuel to produce ..
reformed fuel… which enters on the anode side

$$CH_4 + 2O_2 \rightarrow CO_2 + 2H_2O + e^- + \text{Heat}$$



The process also generates the heat required by the fuel cell

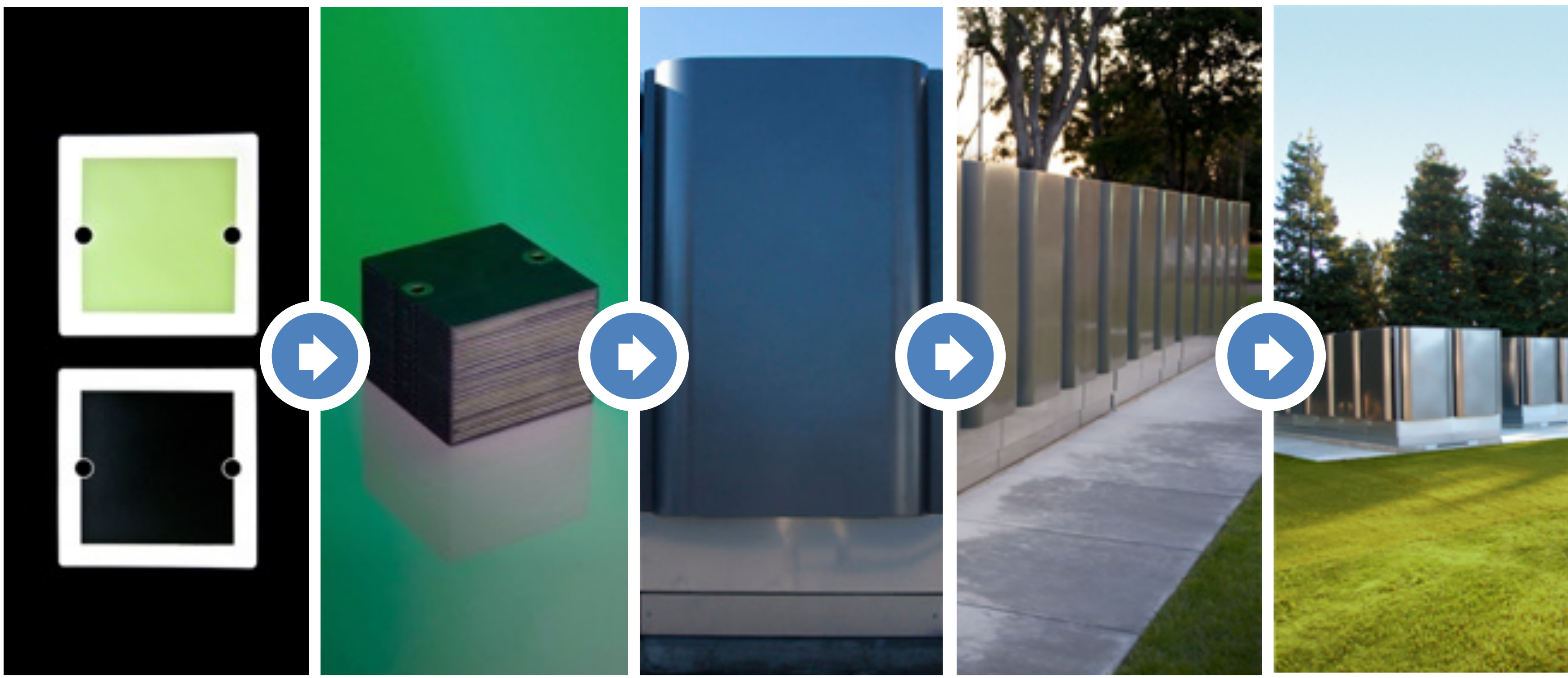


As long as there is fuel, air, and heat, the process continues

Fuel Cell Technology Comparison

Fuel Cell Type	Typical System Size	Efficiency	Applications	Notable US Vendors
Molten carbonate (MCFC)	300kW–3MW	45-50%	Distributed generation, utility	FuelCell Energy
Phosphoric acid (PAFC)	100kW–400kW	40%	Distributed generation	n.a
Polymer electrolyte membrane (PEMFC)	1kW-100kW	35-60%	Backup power, distributed generation, transportation	Plug Power, Altergy
Solid oxide (SOFC)	100 kW–30MW+	52-60%	Distributed generation, utility	Bloom Energy

Scalable, Flexible, Modular



Fuel Cell
25 W

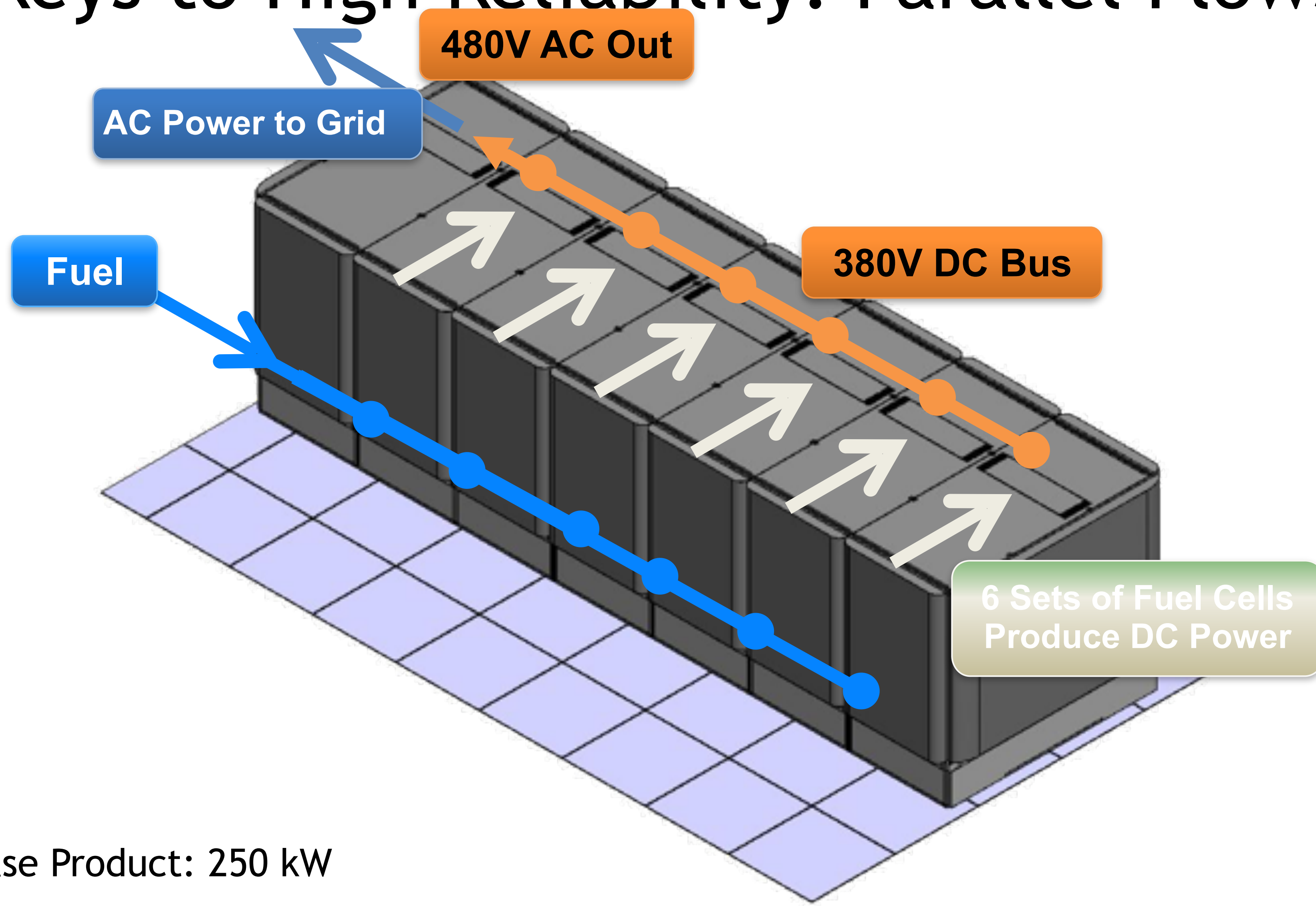
Stack
1 kW

Module
50 kW

System
100 or 250 kW

Solution
100 kW to MW's

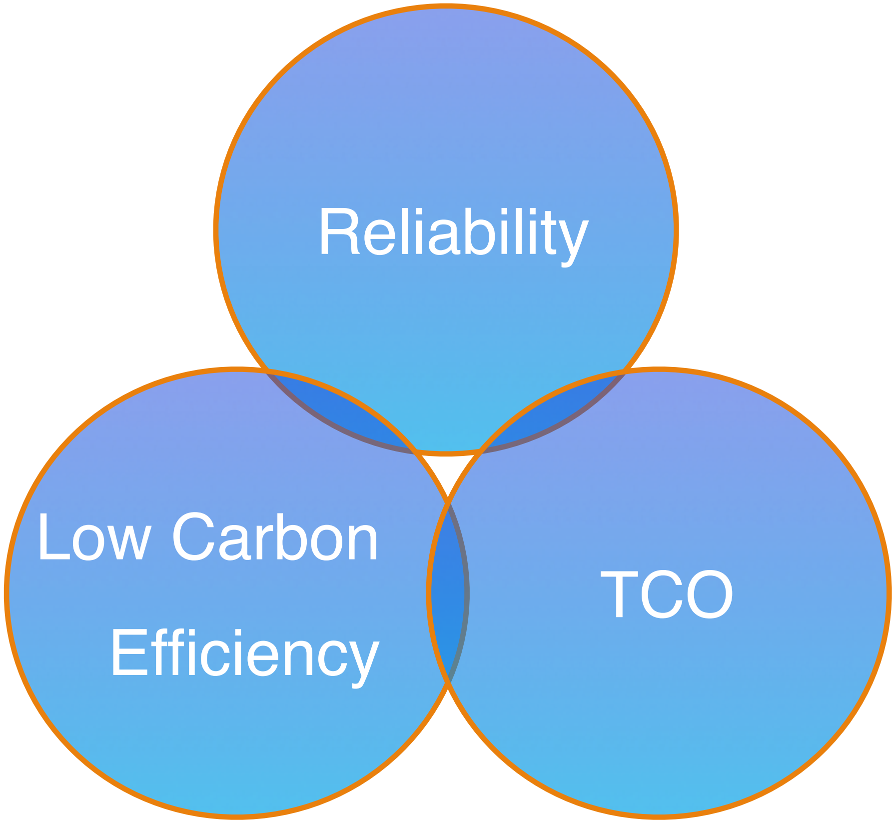
Keys to High Reliability: Parallel Flows



Base Product: 250 kW

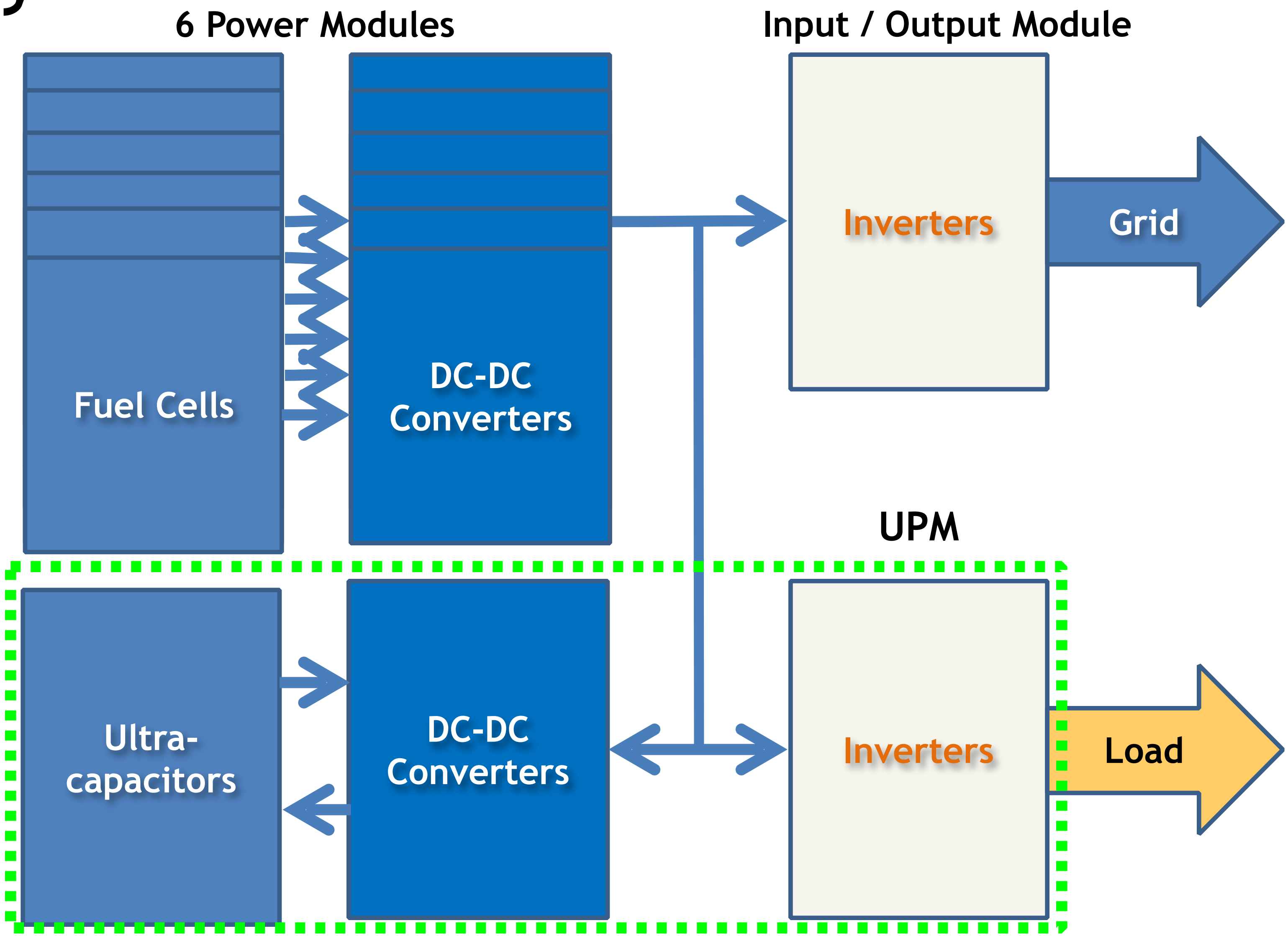
Revolutionizing Data Centers

Data Center priorities-Strike the right balance

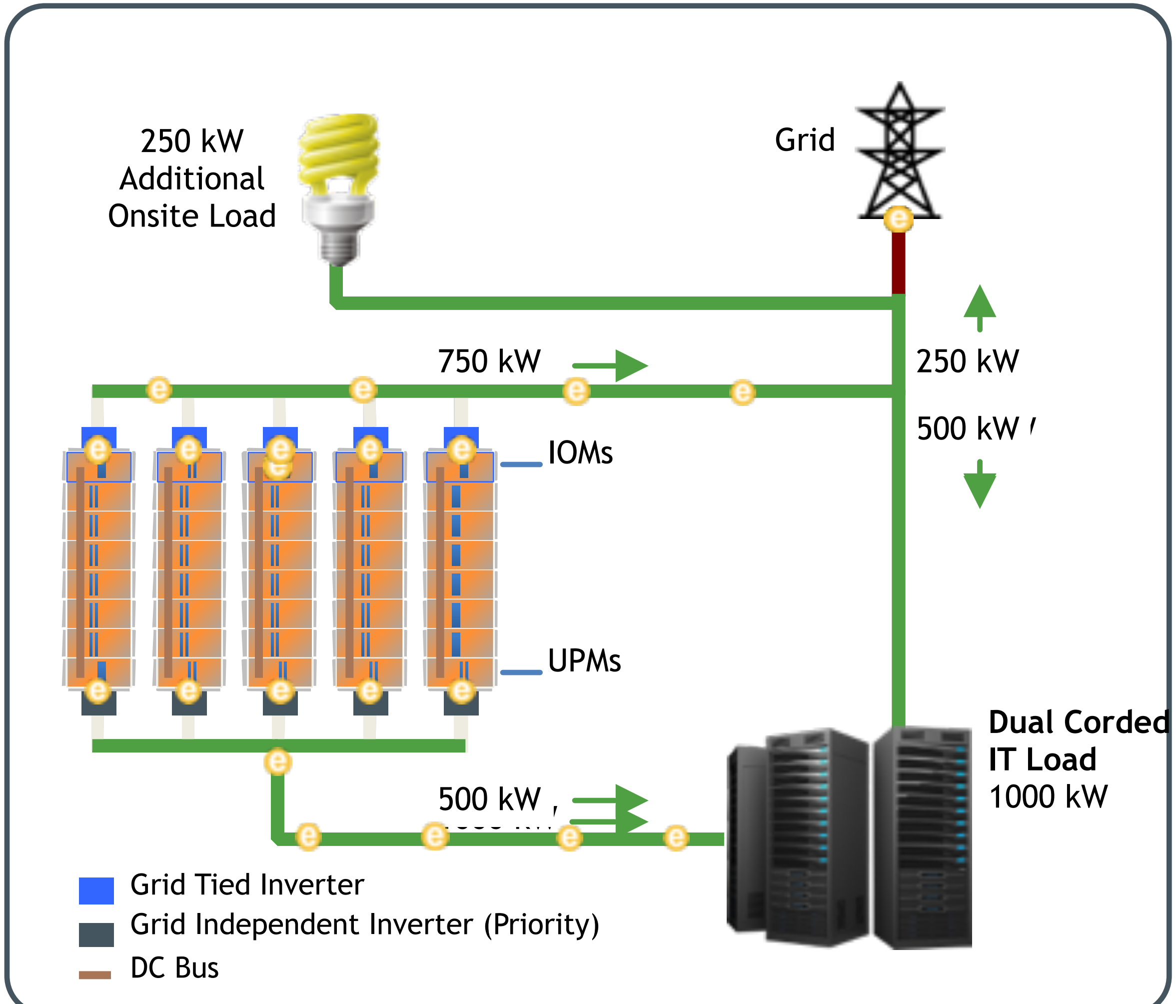


Fuel cells enable the data center to achieve these objectives

System Architecture: Dual Inverters



Mission Critical Solution



- **Better Economics:** displacement of traditional equipment, space reduction, higher efficiency, lower maintenance costs
- **Highest reliability:** two independent sources of power and simplified architecture
- **Gas as primary,** grid as backup
- **Scalability, operational cost predictability, right provisioning**

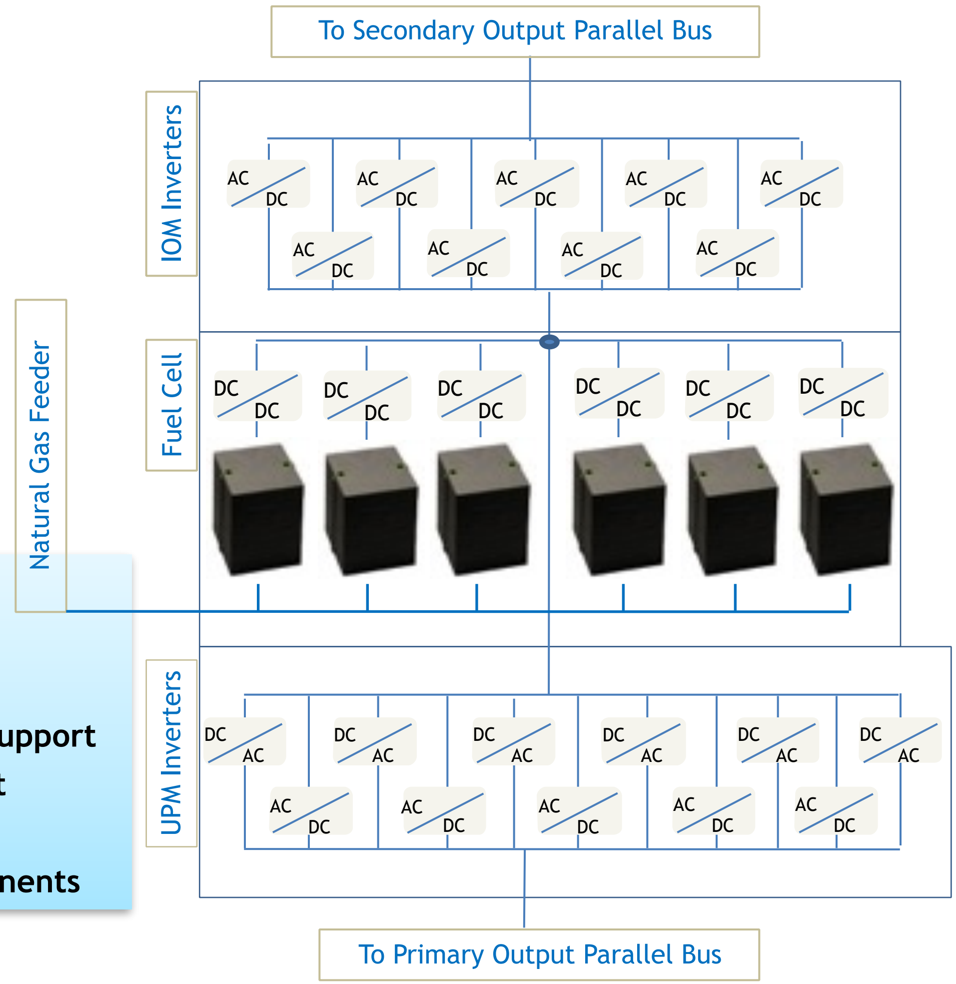
Individual Fuel Cell System (what's in the box?)

KEY DESIGN ELEMENTS

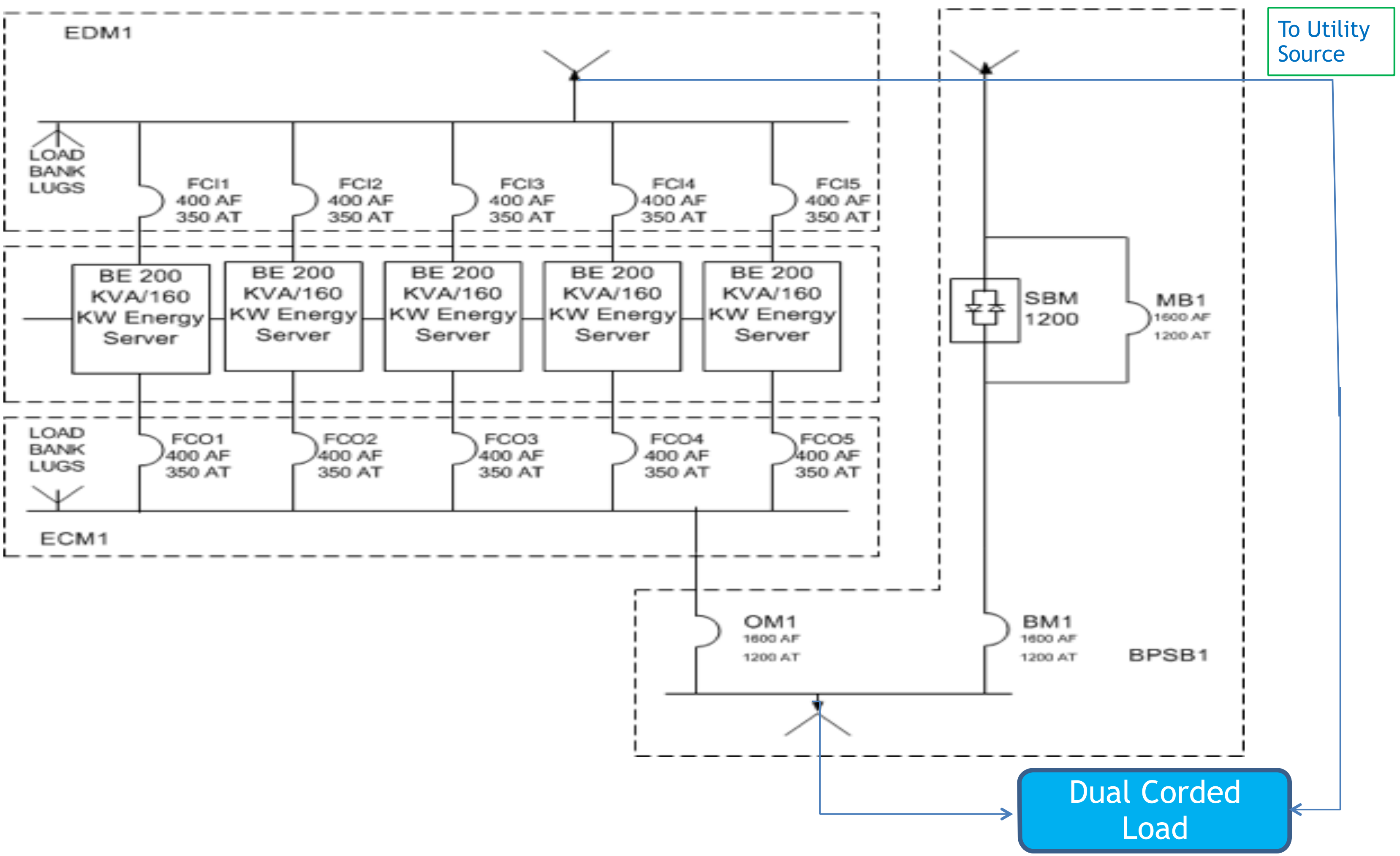
- Redundant Inverters for grid parallel operation
- Redundant inverters for critical output
- Redundant "Hot Box" for energy production
- Modular electronics
- All power electronics are hot swappable

WHY IS THIS IMPORTANT?

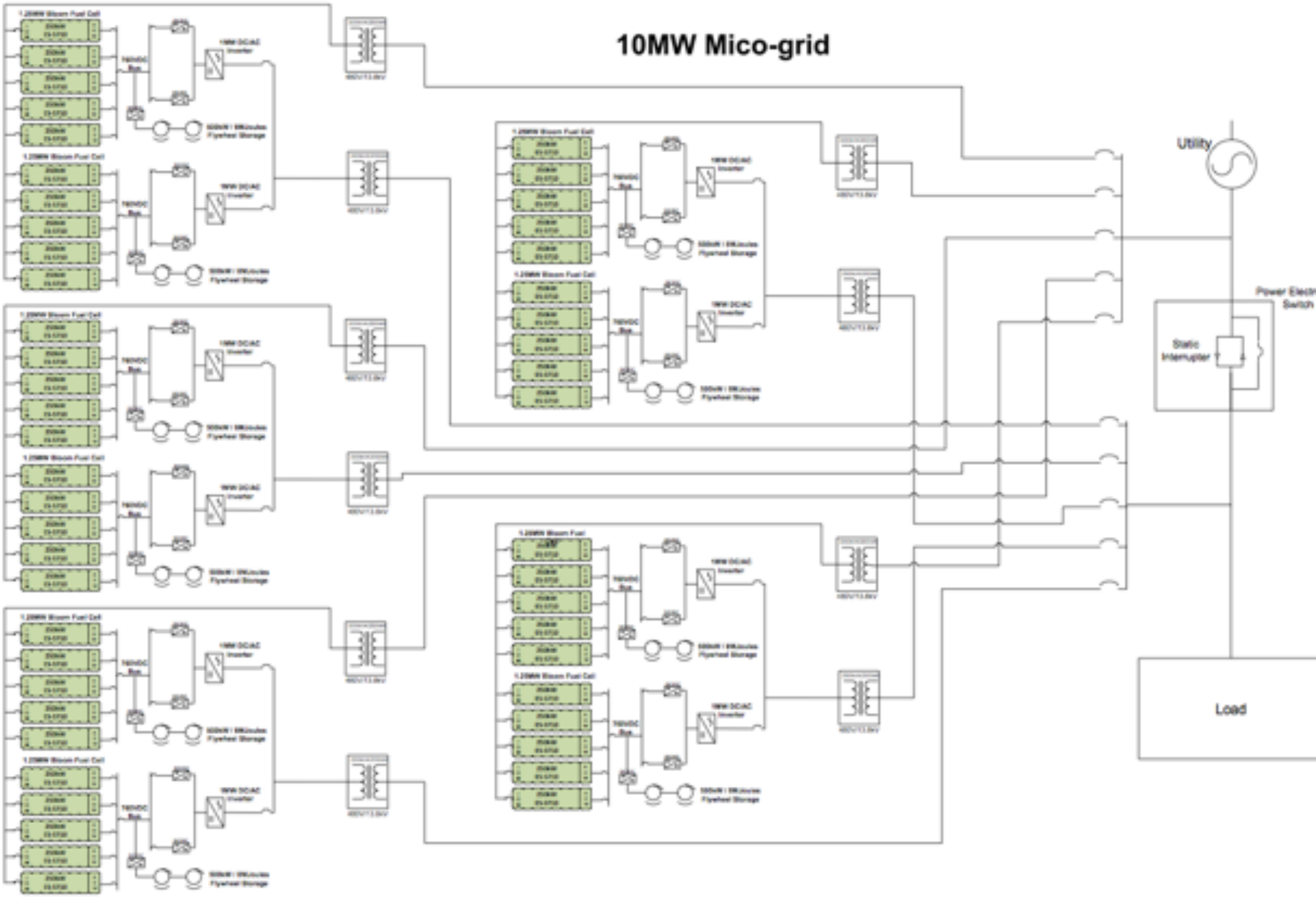
- Key individual system components can be removed for maintenance or repair
- No one power inverter removes total system from bus support
- Control logic ensures all power demand for UPM support supersedes IOM demand
- Redundant hot box for maintainability of critical components



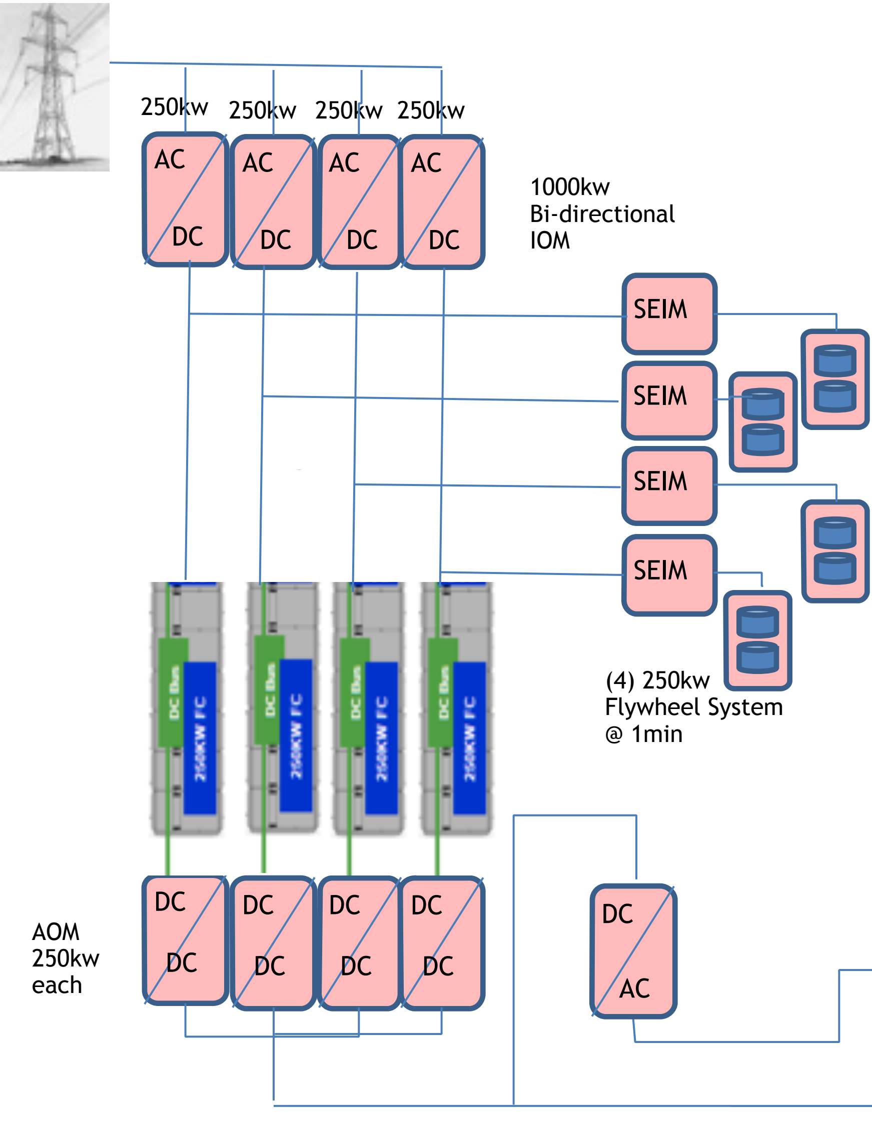
System configuration

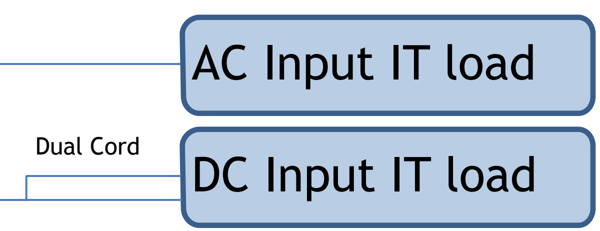


Multi-MW Bus


10MW Mico-grid
Utility
Power Electronics Switch
Static Interrupter
Load

Typical DC Deployment

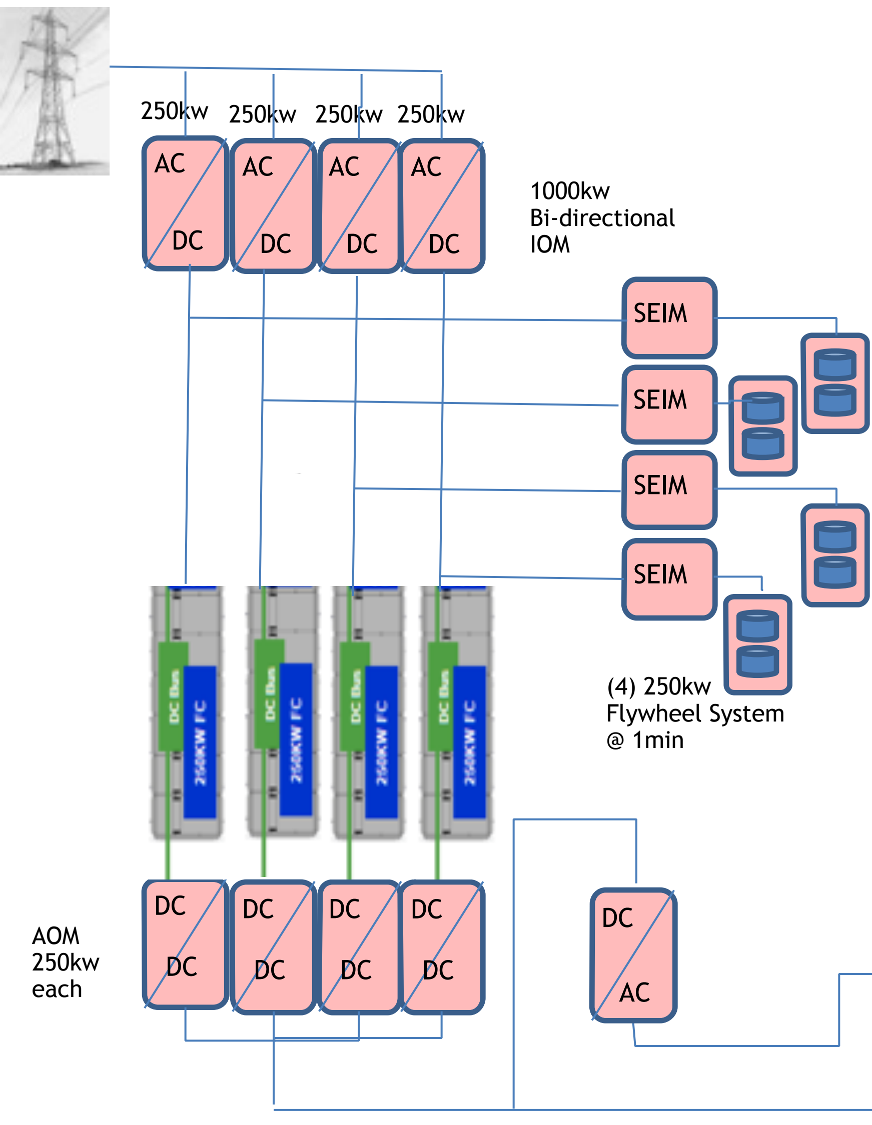




System Features

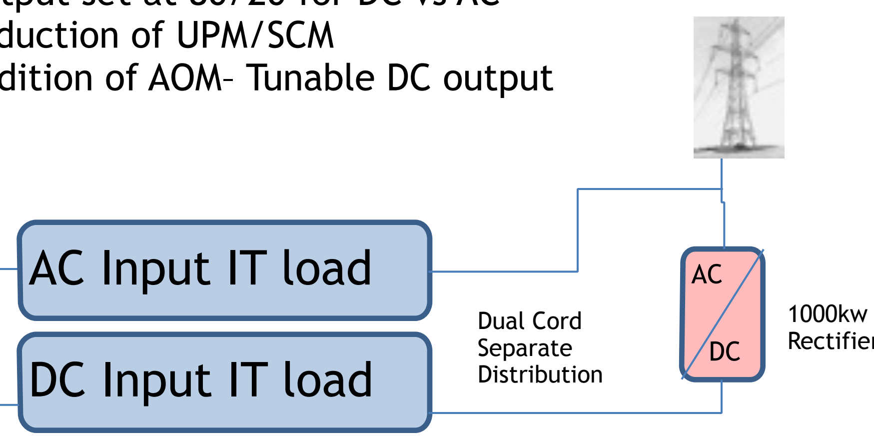
- Redundant Bi-Directional IOM's allows internal distributed bypass
- Fuel Cell used as primary DC generation source
- Total System Design value built via modules
- Common DC Bus configured as +-380vdc
- Stored Energy Interface Modules (SEIM) integrate Flywheel for discharge and charge, sized to handle block loading when in Island mode only
- DC/DC output collector bus produces 380VDC distribution
- DC/AC inverters can be added to serve AC only loads
- All Power Electronic stages can be configured for desired capacity and/or redundancy
- As Shown- 1MW system capacity
- Output set at 80/20 for DC vs AC
- Reduction of UPM/SCM
- Addition of AOM– Tunable DC output

DC Deployment Dual cord



System Features

- Redundant Bi-Directional IOM's allows internal distributed bypass
- Fuel Cell used as primary DC generation source
- Total System Design value built via modules
- Common DC Bus configured as +-380vdc
- Stored Energy Interface Modules (SEIM) integrate Flywheel for discharge and charge, sized to handle block loading when in Island mode only
- DC/DC output collector bus produces 380VDC distribution
- DC/AC inverters can be added to serve AC only loads
- All Power Electronic stages can be configured for desired capacity and/or redundancy
- As Shown- 1MW system capacity
- Output set at 80/20 for DC vs AC
- Reduction of UPM/SCM
- Addition of AOM– Tunable DC output

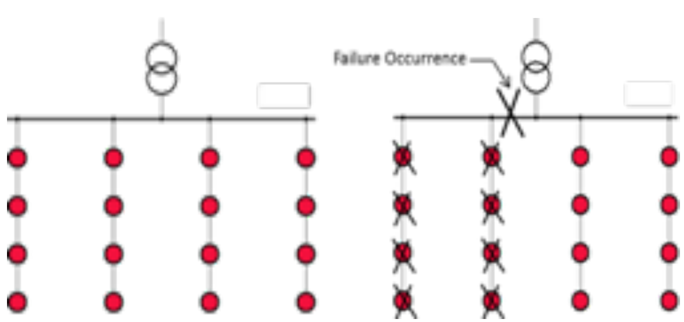


Evaluating System Reliability

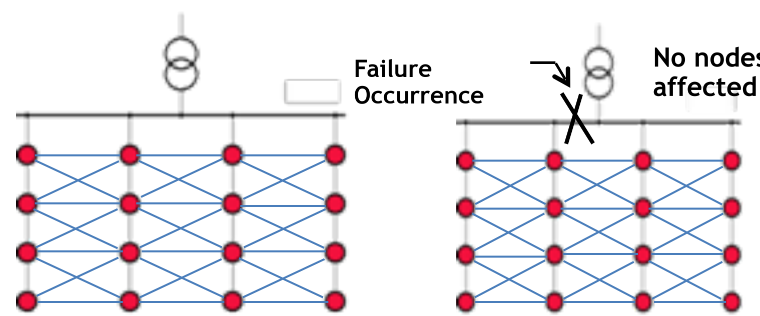
- Reliability evaluation and modeling by outside consultant
- Breakdown the core system components moving from power module features all the way to gas supply
- Probability of failure over a one year mission-0.9%
- Availability of the Bloom system, not including utility (second cord)-99.9977%
- MTBF-994,000 hrs (113.5 years)
- System reliability-99.1%-substantially higher for a dual cord topology

Keys to High Reliability: Gas Infrastructure

	Electric	Gas
Grid Design	***Single Points of Failure***	***Network - Self Healing***
Storage Capacity	None	▪ Long Term – Underground ▪ Short Term – 'Linepack"
Physical Exposure Risk	***High*** *Visible, Above Ground*	***Low*** *Hidden, Underground*
Input Sources	Multiple Diverse set of power producing assets	Multiple Storage, Traditional Gas, Shale Gas, LNG
Reported Availability	***99.9%***	***Not Available*** *Lack of Incidents = Lack of Studies*



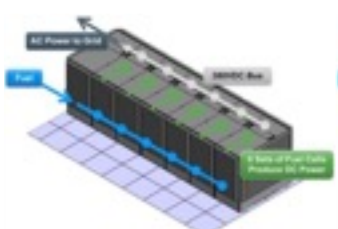
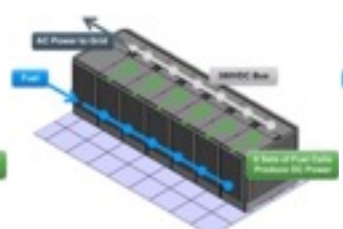
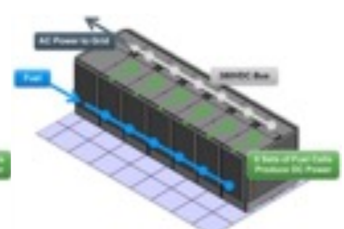
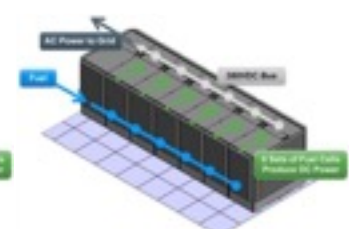
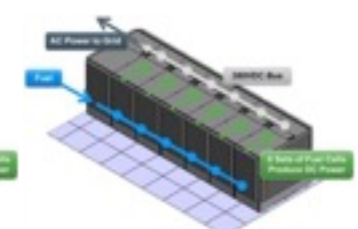
Electrical Grid – Radial Design



Natural Gas Grid – Network Design

Keys to High Reliability

1. Two independent continuously operating sources of power
2. Modularity-Each 1 MW of power utilizes 30 independent power generators operating in concert
3. Redundancy-Dial in your desired topology: N+M
4. Eliminate the traditional points of failure
5. Eliminate complexity and human error
6. Concurrently maintainable
7. Right provisioning-Maximize asset utilization by feeding excess energy to additional onsite loads
8. Comprehensive remote monitoring
9. Mature/proven technology

Bloomenergy®

Be the solution

Is N Enough?

DIGITAL REALTY

Jim Smith – CTO
1 June 2015

Senior management question:

What is the right balance between expenditure and performance for our business?

First Principles
Reliability vs Availability

THEY ARE NOT THE SAME

- Reliability is the probability of failure over a discrete time period
  - Usually a single number, expressed as a % or equivalent
- Availability is the “usability” of a service or feature over a continuous period
  - Enter a compound equation adding time-to-repair and time-to-restore
- Our systems are a complex blend of these basic concepts

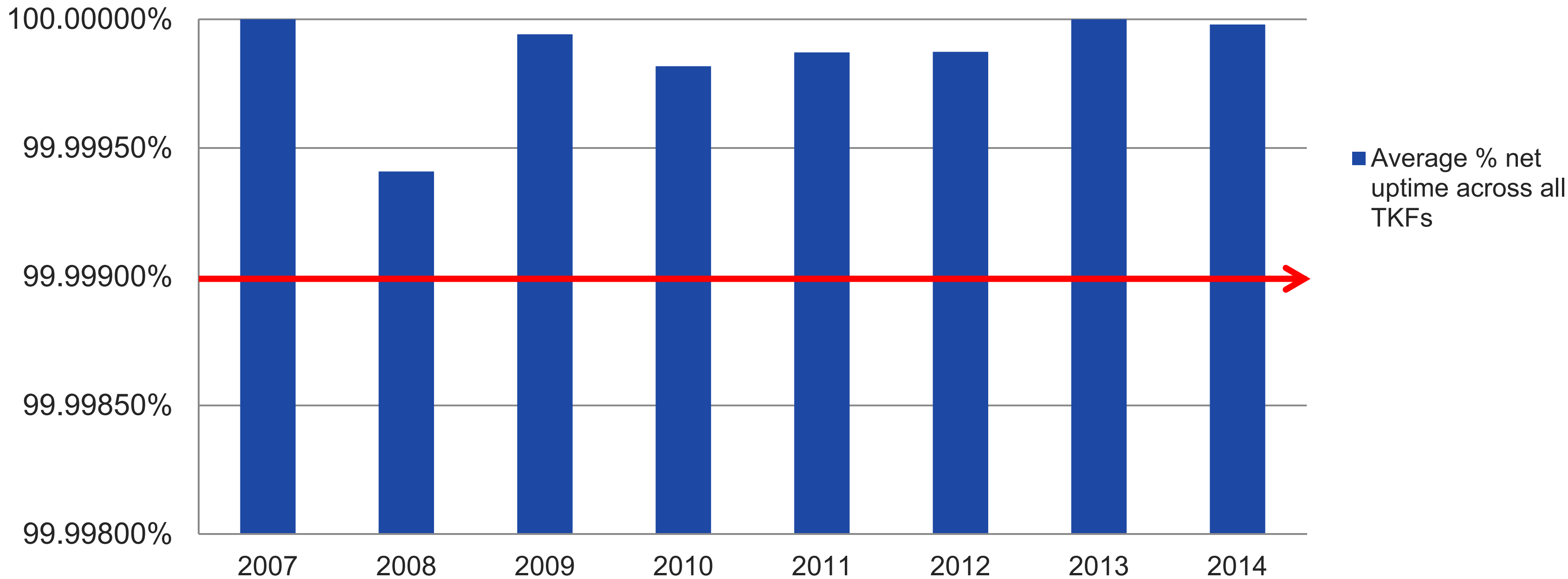
Qualitative Insights (Positive Mythologies?)

YOUR ENGINEERING INTUITION IS JUST AS VALUABLE AS THE MATH

- There is a balance point between increasing complexity and overall reliability/availability
- Time to restore will/should end up driving your decisions
  - Automated, fast time to restore can tolerate lower reliability and still drive high availability
  - Complex, long time to restore may demand higher reliability in order to ensure high availability
  - Positive industry trends right now – "cattle not pets" concepts, software layer fault tolerance, orchestration layers for vm provisioning, etc.
- "How things fail" matters greatly
- Invest in reliability closer to the load
- Actual availability distributions are lognormal/fat-tailed
  - HUMANS are involved
  - Software too….

Availability Analysis – Summary

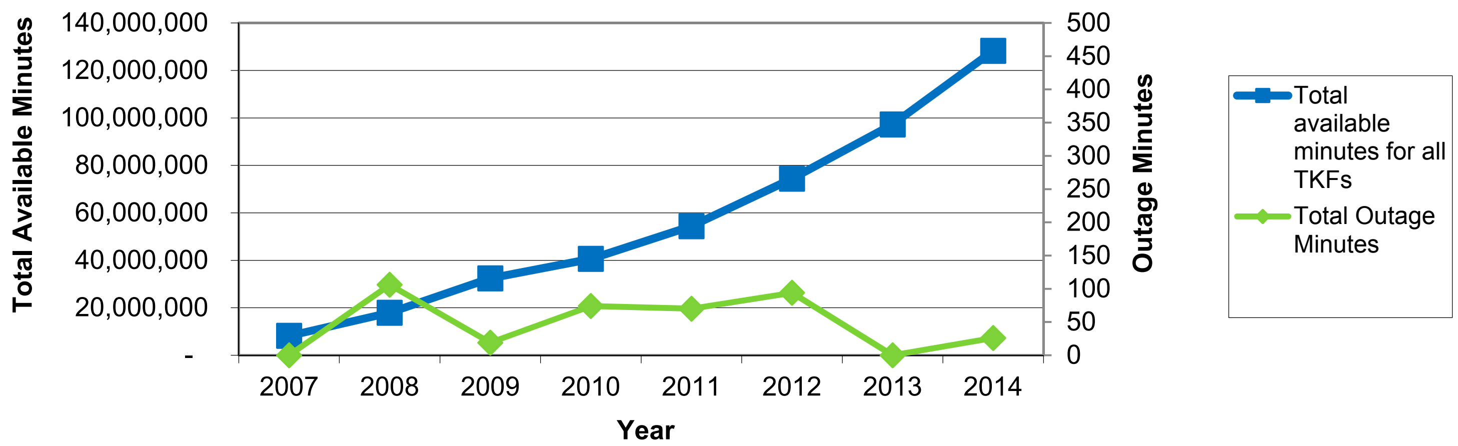
Turn-Key Flex[SM] Data Centers Availability



1. Data collected from all Turn-Key Flex[SM] data centers designed and operated by Digital Realty.

Availability Analysis

Turn-Key Flex℠ Data Centers – Available Minutes

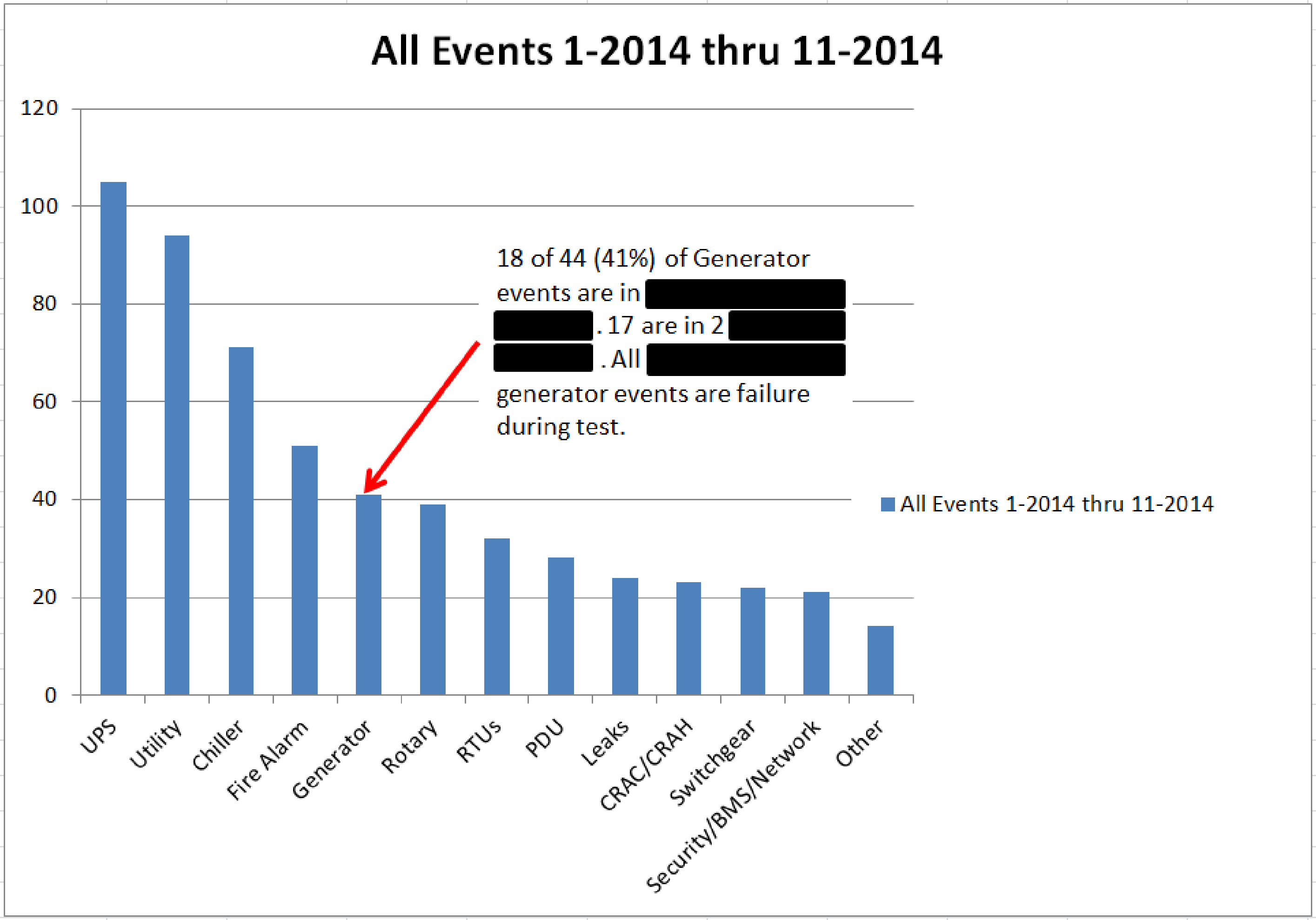


"Total Outage Minutes – Service impacting" (left axis) has decreased or remained flat while the number of "Total Available minutes" (right axis) has increased over 1,469%.

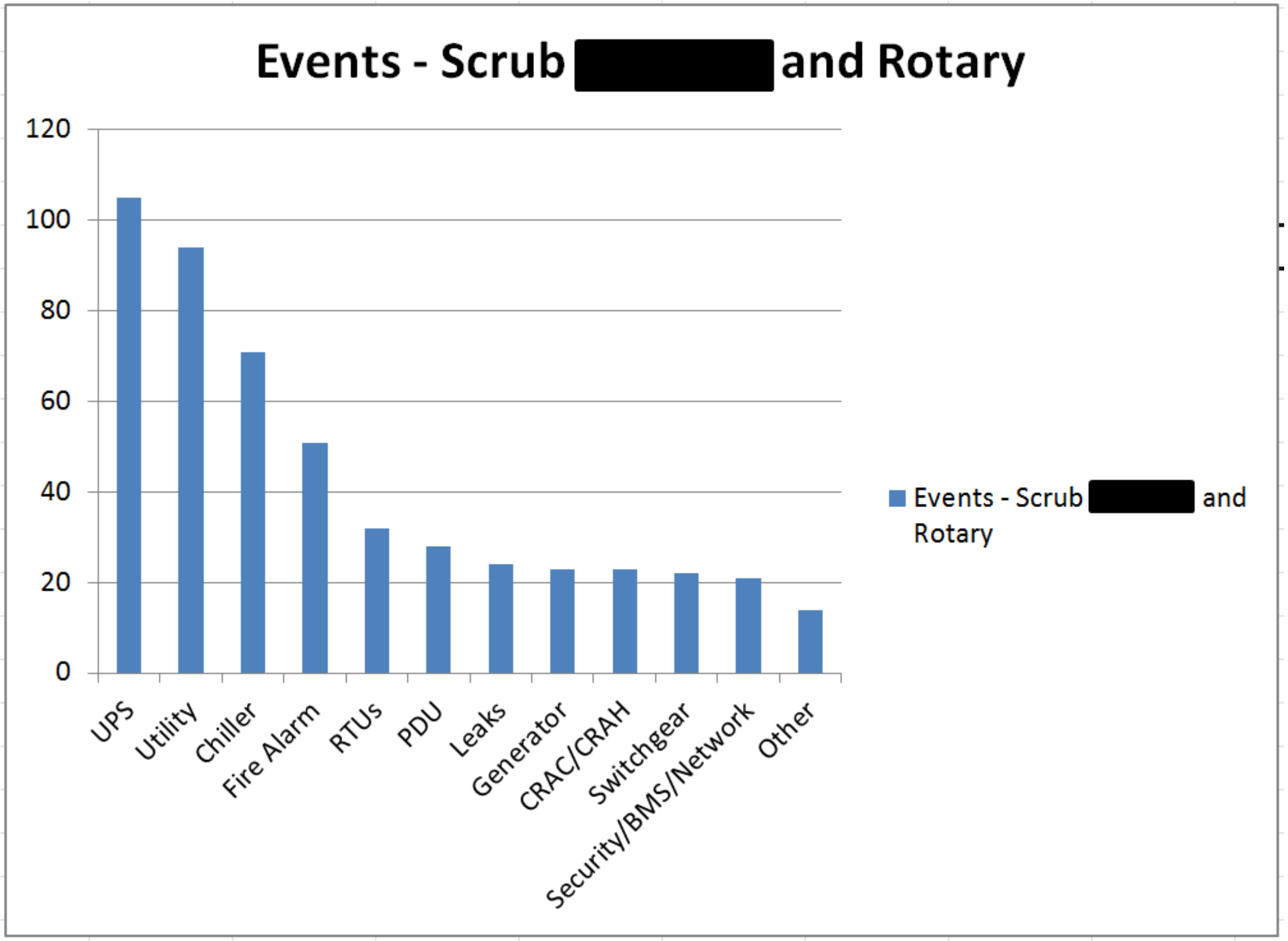
- *Over 453 million operational minutes*[1]
- *863 site/years of operational experience*[2]

1. Total operational minutes prorated based on operational start date per data center for the period 2007 – 2014.
2. Site years equals total operational minutes during period 2007 - 2014 divided by number of operational minutes per year per data center (or 525,600 minutes).

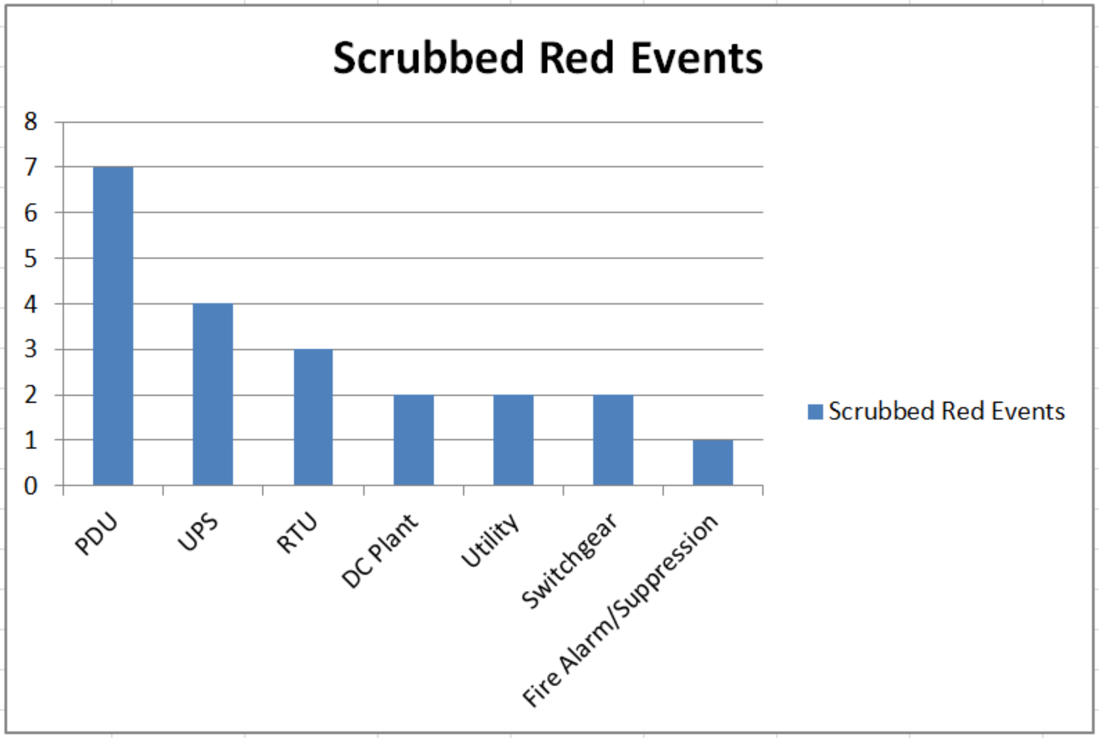
Event Parse 2014



Event Parse 2014



Event Parse 2014



Supports traditional and academic theories that failures occurring closer to the load are more impactful/RED

Soft Conclusions

- "Events" impact customers but may not result in impact to loads
  - Fire alarms, security access, loading dock, etc.
- Red events bifurcate into two sample sets:
  - Compound failure modes, usually during maintenance activities. We have seen three failures stacked…
    - Larger impact, bigger system failures
  - One person moving breakers in the PDU/RPP infrastructure
    - "Contained" impact, easier to recover
- Change Management and people management are by far the most important feature of high availability
  - Was not always the case
  - Our industry should be proud of the professionalism we have brought to customers and workloads

Recommendations/Challenges

- Automated, fast restoring systems can benefit the most from reduced resiliency
  - But, smaller (<1MW) datacenters and distributions can get quite expensive to multi-tier
- Harder to restore systems may benefit from “traditional” high availability methods and systems
  - Enterprise complexity
- What to leave out?
- What is the “cost” of variability/multi-tiering in your fleet?

DIGITAL REALTY

Jim Smith
CTO, SVP of Operations
jsmith@digitalrealty.com
www.digitalrealty.com

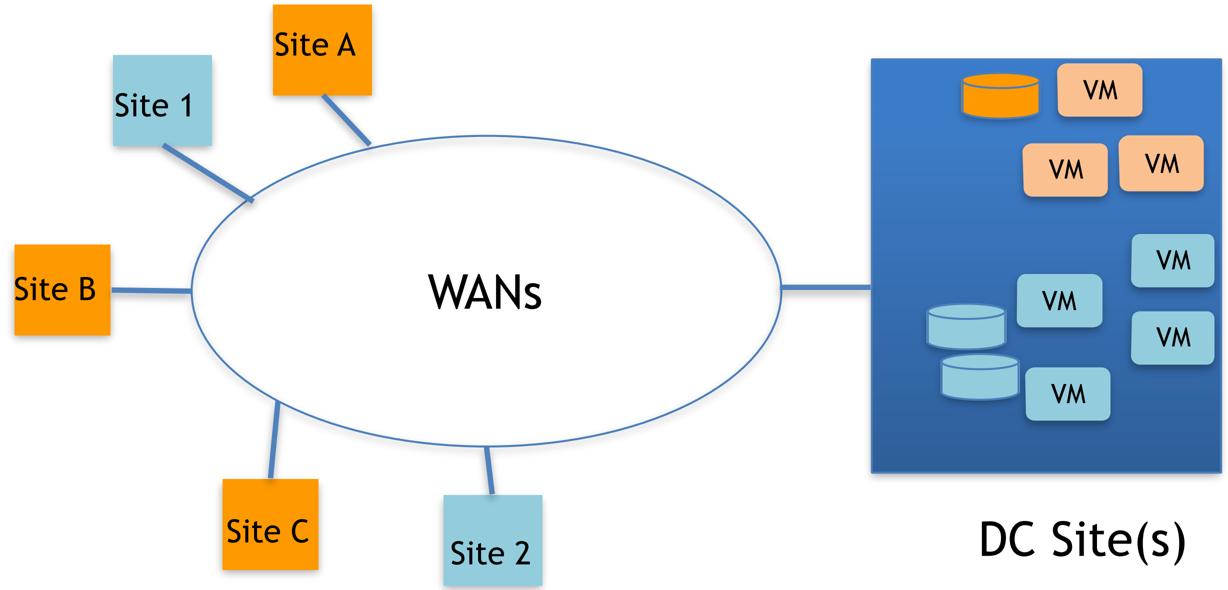
Inter-AS Options for NVO3 and BGP/MPLS VPN Interworking

Lucy Yong
lucy.yong@huawei.com

June 2015   San Francisco USA

Datacenter Cloud Services for Enterprise

- An enterprise (orange) buys a cloud service from DC SP:
  - Cloud services only are accessible at Enterprise sites
  - DC SP uses many VMs to run the cloud service at the site(s)
- An enterprise (blue) buys computing resources from DC SP:
  - Run its applications on the VMs/storages at DC site
  - Access the applications from Enterprise sites
- Both cases require service privacy for Enterprise customer access only
  - Virtual private network is necessary in DC and WAN for the privacy



NVO3 in Data Center

- Network Virtualization Overlay (NVO3):
  - Driven by server virtualization paradigm,
  - Decouples tenant hosts from physical network,
    - Tenant hosts are not constrained by physical network subnets,
    - Allows separated tenant address spaces,
    - Supports many tenant virtual networks over one physical network
    - VMs can move from one server to another w/o reconfiguration.
- SDN controller is for NVO3 control plane:
  - A centralized control plane,
  - Decouples control plane and data plane,
  - Binds virtual network (VN) layer and VN transport layer via the mapping,
  - Different from traditional BGP/MPLS VPN.
- Data plane couples VN layer and VN transport layer:
  - Forwarding based on tenant end point/tunnel end point mapping,
  - Tenant packets are encapsulated, delivered via IP tunnels,
  - Different from BGP/MPLS VPN.

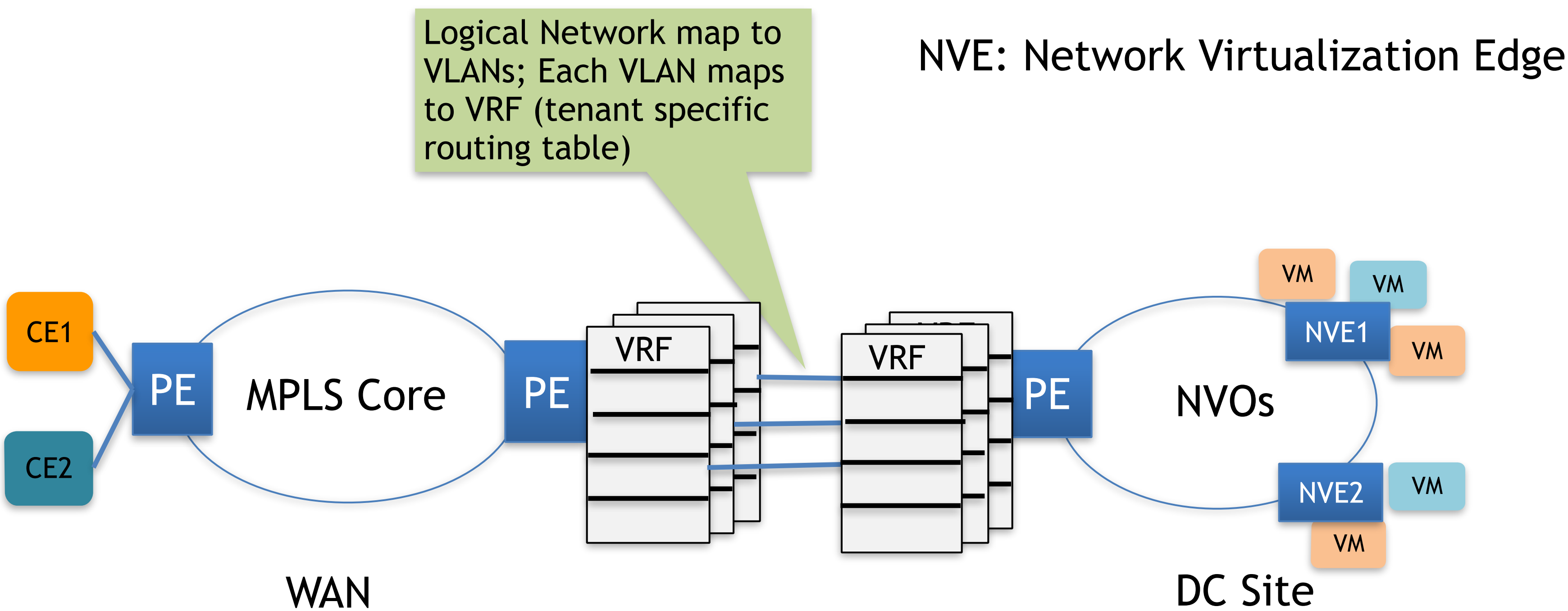
Motivation on use of inter-AS

- Tenant networks in a DC need to communicate with tenant sites outside DC in private /secure manner
- BGP/MPLS VPN inter-AS is widely deployed and supported by many service providers
- Tenant sites already used BGP/MPLS VPN over WAN
- DC and WAN Service Providers often belong to different administrations, i.e., different AS
- NVO3 and BGP/MPLS VPN have some common properties and can interconnect each other
  - VPN technique
  - use of tunnel technology to delivery VN packets

Leveraging existing VPN inter-AS is desirable

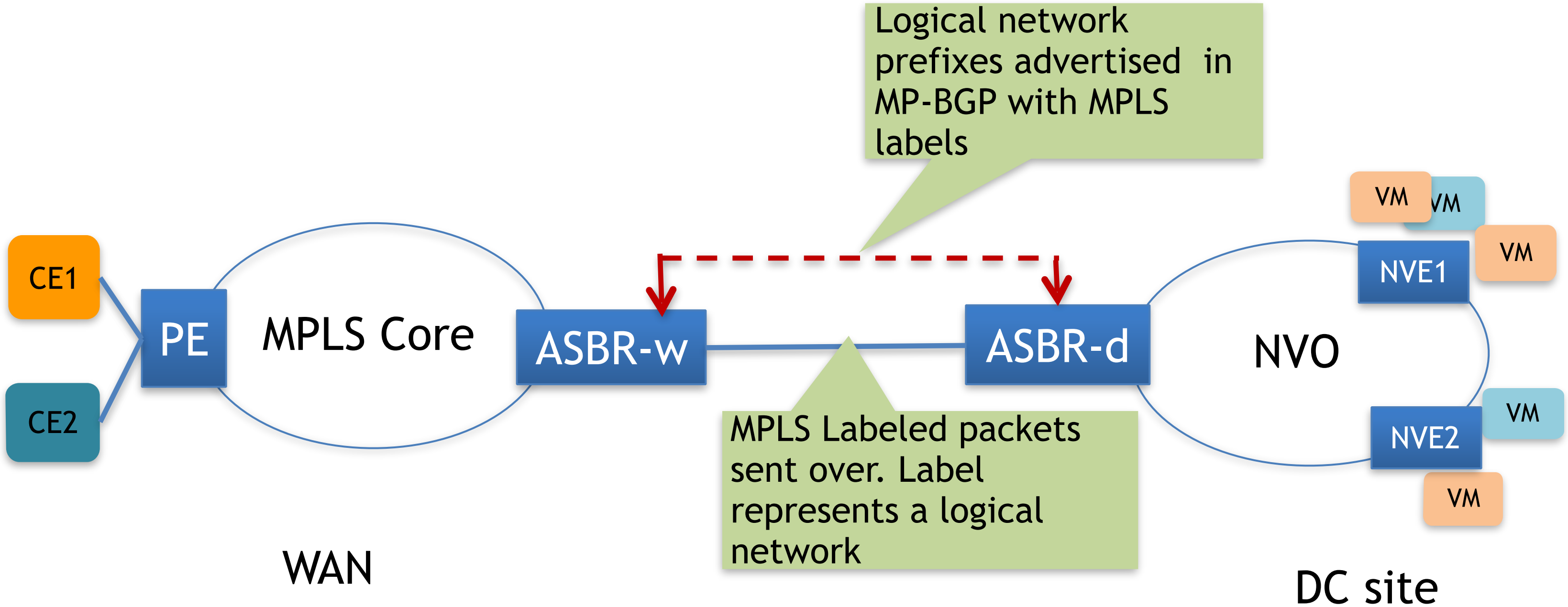
Option A for NVO3 and BGP VPN

- Inter-AS Option A [RFC4364]: map VNs to VLANs.
- WAN and DC VPN implementations are independent



Option B for NVO3 and BGP VPN

- Inter-AS Option B [RFC4364]: support MPLS b/w ASBRs
  - No sub-interfaces and VRF lookup at ASBRs

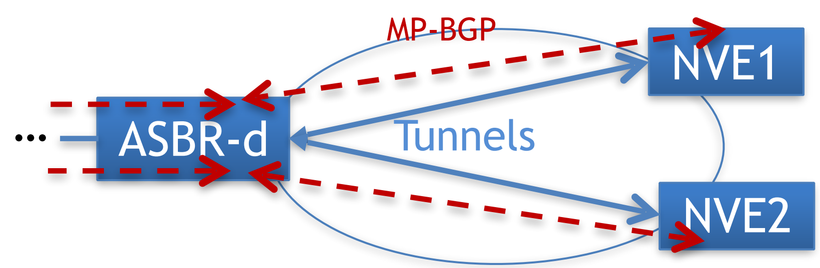


Problem: DC site uses IP tunnel while WAN uses LSP path

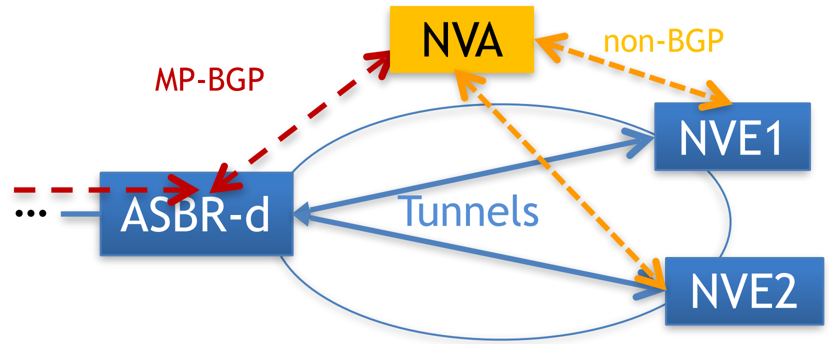
Require: ASBR-d has to stitch IP tunnel and LSP path

Option B Implementations at DC site

- Case 1: Vanilla Option B by use of BGP in DC.
- Case 2: Vanilla Option B by use of NVA in DC.
  - NVA: Network Virtualization Authority for NV03 control plane
  - NVE: Network Virtualization Edge for NV03 data plane functions



Case 1



Case 2

- Case 1 & 2 are in common at following:
  - ASBR-d does not have VRFs and sub-interfaces
  - ASBR-d has the exact same tunnel stitching behavior

ASBR-d Operation in Case 1&2

Control Plane:

- Allocates a VNID per a VPN label when getting a route/label from ASBR-w.
  - VNID is Virtual Network Identifier
- Allocates an MPLS label per <NVE & VNID> when getting a route from NVE.

Data Plane:

- Has an outgoing forwarding table,
  - An entry has the mapping of VNID to MPLS label.
- Has an incoming forwarding table,
  - An entry has the mapping of MPLS label to < NVE IP address, VNID>.
- Performs tunnel stitching between IP/VNID and MPLS label,
  - No payload lookup at ASBR-d.
- No change to ASBR-w

VNID	Out VPN Label
10000	3000
10001	4000

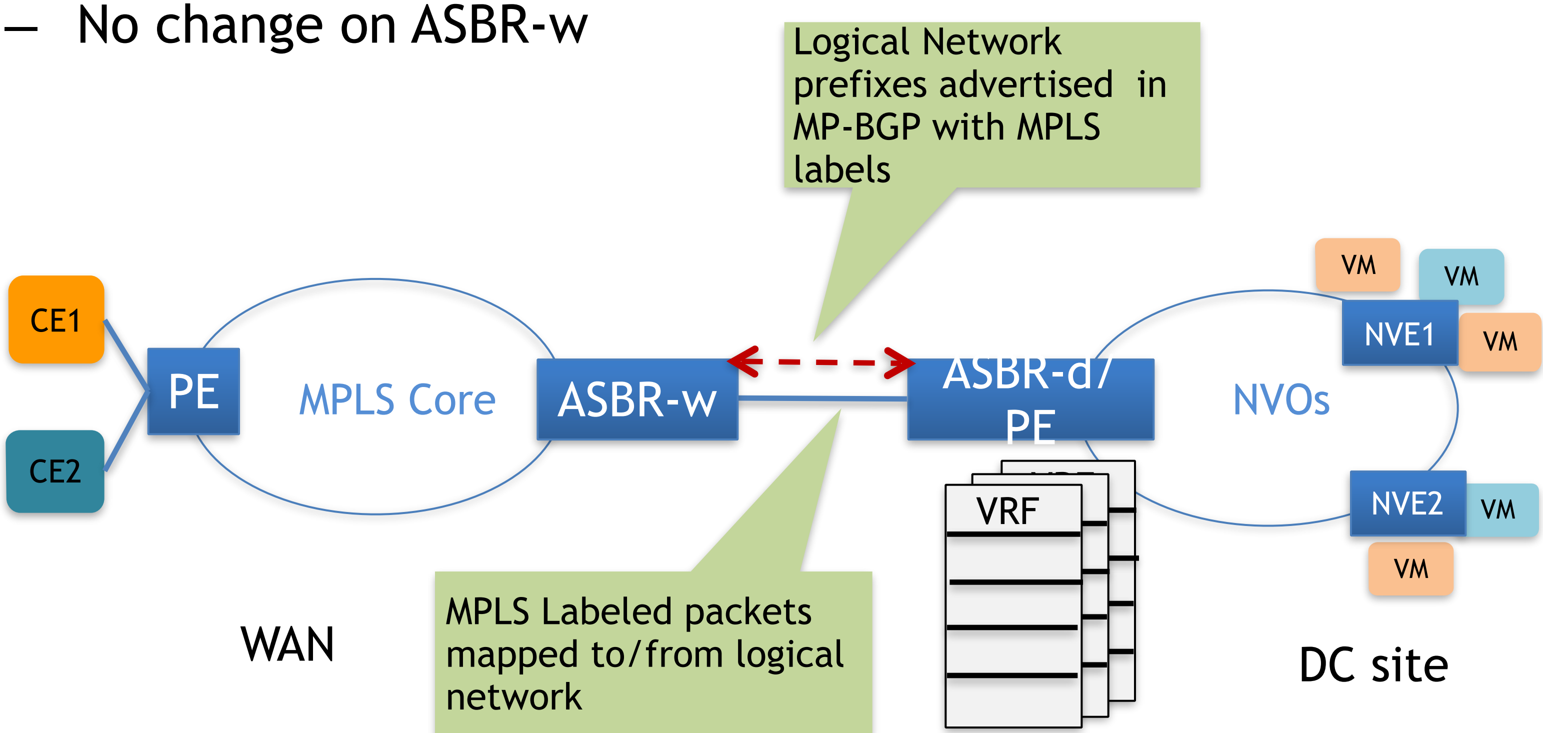
Table 3 Outgoing forwarding table

MPLS VPN Label	NVE + VNID
1000	NVE1 + 10
2000	NVE1 + 20
1001	NVE2 + 10
2001	NVE2 + 20

Table 1 Incoming forwarding table

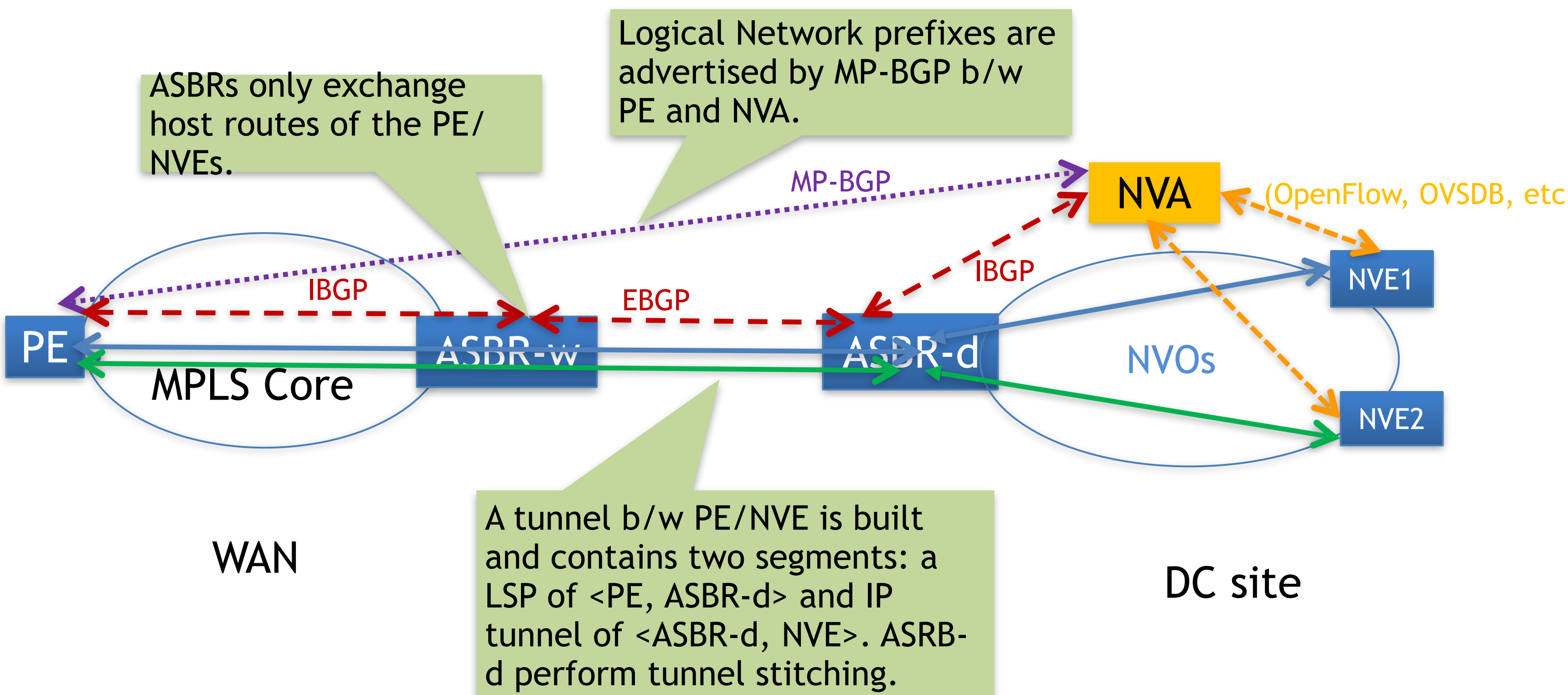
Option B Implementations at DC site

- Case 3: Partial Option B.
  - No sub-interface b/w ASBRs,
  - ASBR-d performs VRF lookup, instead of tunnel stitching
  - No change on ASBR-w



Option C for NVO3 and BGP VPN

- Two Cases: BGP on NVE or BGP on NVA
- Multi-hop EBGP is used to distribute PE/NVE routes b/w PEs and NVA (NVE)
- ASBRs exchange PE/NVE routes and are not aware of VPNs
- PE and NVA (NVE) uses MP-BGP to exchange labeled VPN routes



ASBR-d for Option C

- Control Plane:
  - WAN -> DC: ASBR-d allocates IP address per PE
  - DC -> WAN: ASBR-d allocate label per NVE
- Data Plane:
  - Stitches IP tunnel and LSP tunnel
- Operation Rule:
  - Partitions VNID space: lower 1M for inter DC (to match label size), rest for intra DC
- BGP extension
  - A new SAFI for the NLRI that contains NVO3 tunnel info: tunnel encap. type, tunnel EP address

Options Comparison

- Inter-AS options are good for the cases where DC and BGP/MPLS WAN belong different administrations
- ASBR-d implementations are different in options

Items	Option A	Vanilla Opt-B	Partial Opt-B	Option C
Sub-	*Yes*	*No*	*No*	*No*
VRF	*Yes*	*No*	*Yes*	*No*
Configuratio	*Complex*	*Simple*	*Moderate*	*Simple*
Scalability	*Worst*	*Very Good*	*Better than opt-*	*Best*
HW Support	*Support*	*Upgrade*	*Support*	*Upgrade*
Security	*Best*	*Weak*	*Moderate*	*Worst*
BGP/MP-BGP	*Support*	*Upgrade*	*Support*	*upgrade*
Complexity	*Low*	*Middle*	*Middle*	*High*

Q&A

- Operator feedbacks on these approaches:
  - Which option fits your environment?
  - Operational concerns and experience?
  - Additional necessary techniques?

- For detail:

draft-hao-bess-inter-nvo3-vpn,

draft-hao-bess-inter-nvo3-vpn-optionc.

Open-IX Update
NANOG 64

OPEN-IX

Update Points

- Certification Status
- Data Center Evaluation Tools
- Transparency/Complaint Resolution
- Standards Expansion/Refresh
- Discussion

OIX-1 Certification (5 IXPs)

- Completed: Four (4) MSAs, Five (5) IXPs
- Pending: One (1) new MSAs, One (1) IXP

COMPLETED (5)

Amsterdam Internet Exchange
- NY/NJ
- SF Bay Area
- Amsterdam

London Internet Exchange
- VA/MD/DC

Deutscher Commercial Internet Exchange
- New York

PENDING (1)

Florida Internet Exchange
- Miami, FL

ANNOUNCED (3)

Omaha Internet Exchange
- Omaha, NE

Phoenix NAP
- Phoenix

Amsterdam Internet Exchange
- Chicago

OIX-2 Certification (24 DCs)

- Completed: 22 Data Centers, Ten (12) Companies
- Pending: Five (5) Data Centers, Four (4) new Companies

COMPLETED (22)

CyrusOne

1. Dallas, TX
2. Houston, TX
3. Austin, TX
4. Cincinnati, OH
5. Cincinnati, OH
6. Phoenix, AZ

DIGITAL REALTY

1. Los Angeles, CA
2. San Francisco, CA
3. Dallas, TX
4. New York, NY

QTS

1. Atlanta, GA
2. Suwanee, GA
3. Richmond, VA

1. Ashburn, VA
2. Piscataway, NJ

SENTINEL

1. Durham, NC
2. Somerset, NJ

zayo

1. Atlanta, GA

CONTINUUM DATA CENTERS

1. Chicago, IL

evoswitch

1. Manassas, VA

edgeconnex

1. Houston, TX

VANTAGE DATA CENTERS

1. Santa Clara, CA

DATAGRYD

1. New York, NY

SABEY Data Centers

1. New York, NY

PENDING (5)

zayo

1. Miami, FL

phoenixNAP GLOBAL IT SERVICES

1. Phoenix, AZ

DATA BANK

1. Richardson, TX

experis

1. Middletown, VA

Jaguar network

1. Marseille, FRANCE

Data Center Evaluation Tools

- Questionnaires – Go beyond design and operational compliance
- Technical and Non-Technical Aspects
  - Communication factors
  - Service factors
  - Turn-up Factors
  - Access factors
  - Networking factors
  - Energy factors
  - Human factors
  - Certification factors
- Different Format/Use
  - Generic
  - Open-IX Branded
  - Co-Branded

Transparency/Complaint Resolution

- A complaint link will be added to the website, with an explanation of the process
- Complaints go to the appropriate committee head, and the board for reference
- Committee head or a designee will reach out to both parties within one week to discuss the complaint
- Committee head or a designee will work with the certified entity to try to reach resolution within two additional weeks
- Certified entity will then have up to four weeks to implement remedial action
- If resolution cannot be reached or certified entity does not complete remedial effort, certified entity shall first have the right to forfeit their certification
- Results of the process shall be reported to the board within eight weeks
- If the violation remains and the certified party will not voluntarily forfeit, the board will commence a formal process to withdraw certification, which will include comment on the members list

Standards Expansion/Refresh

- Single-corded line-up/reduced reliability
- Virtual peering and variants thereof
- Remote operated/lights out facilities
- Strengthen core standards
- Continue to add issues that impact a significant number of the constituents

Discussion

OIXA
CERTIFIED