Long Distance Disruption-Tolerant Wireless Networking

Kevin Fall PhD
Intel Research, Berkeley & WHOI ["soon"]
kfall@intel.com
http://WWW.DTNRG.ORG

Nov 9, 2005 / RVTEC / Oregon State Univ.

Who am I?

- Intel Research (Berkeley)
  - Part of Intel Corporation (the chip people)
  - One of the ‘lablets’ working with universities
- Chair of the ‘Delay Tolerant Networking’ RG
  - Part of IRTF/IETF (the people who do standards)
  - Participant/PI in related NSF and DARPA programs
- I am here because…
  - Our work might help oceanography
  - I am interested in [some of] your problems

802.11 Standard

- Half-duplex radios
  - only one of: Transmit or Receive
  - designed for low-delay office environment
- Collision avoidance
  - Cannot listen to channel while transmit
  - Backoff before every transmit
  - Contention for access to broadcast channel

802.11 Packet Sequence

Packet transmission time : Tp
ACK transmission time: Ta
Propagation delay : Td
ACK Timeout > (Tp + 2*Td + Ta)

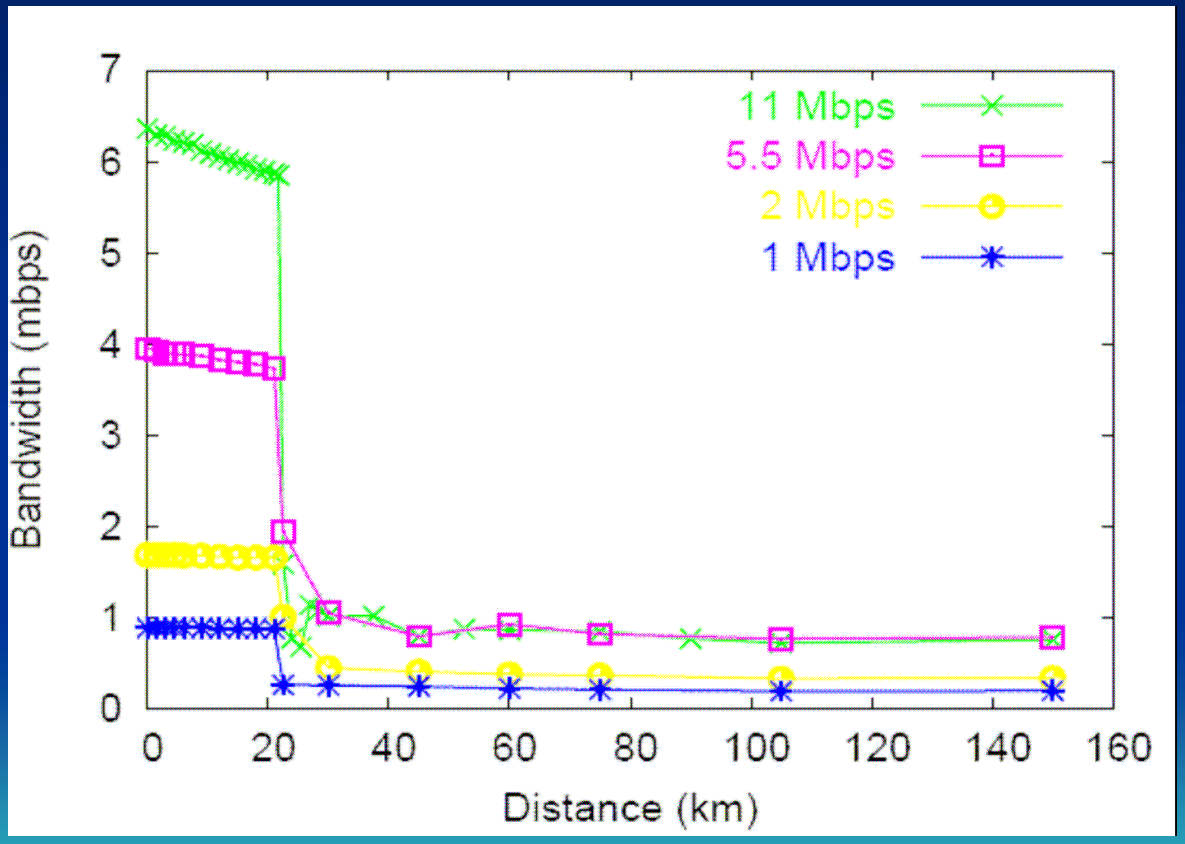
Features:

- Stop and go
- MAC retransmissions for reliability

Key Issues:

- Channel utilisation
- ACK timeout
- Collisions

Throughput vs distance



- PRISM 2.5 Radios
- UDP throughput
- With retries enabled

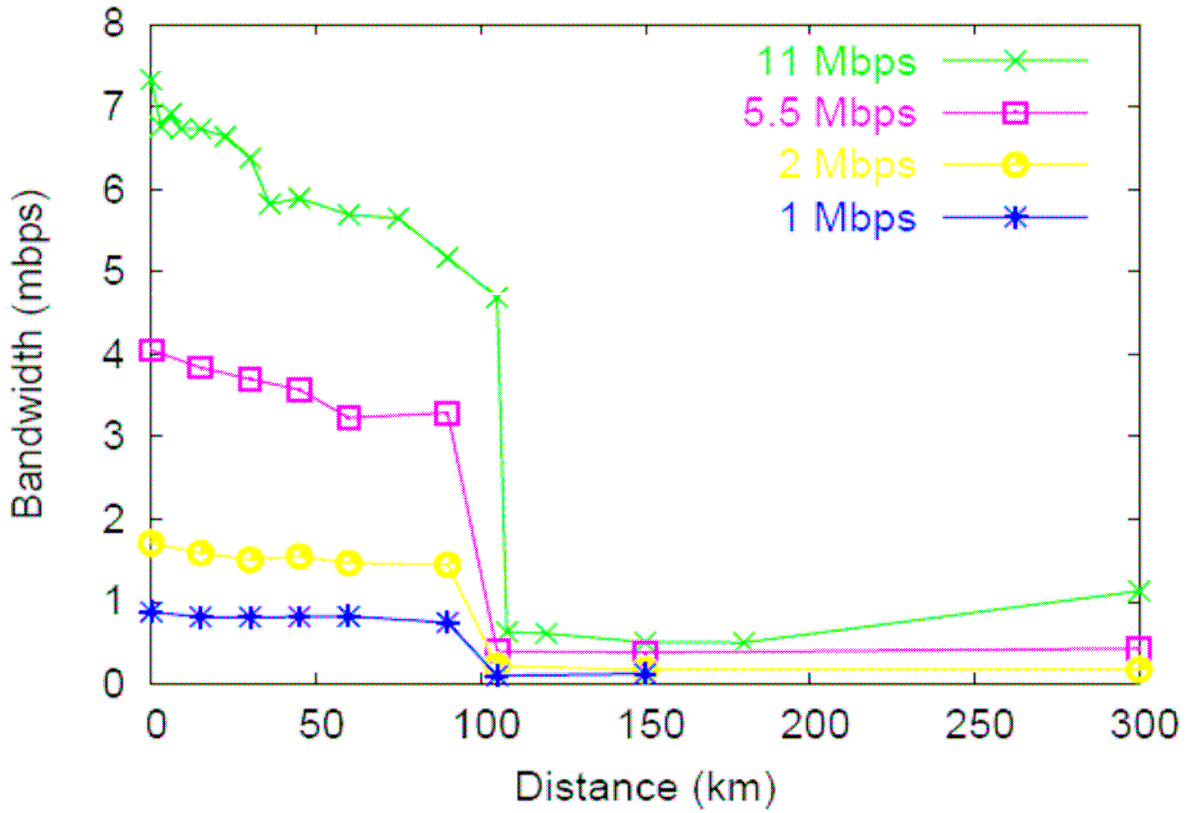
Results:

- Gradual drop as round trip time increases
- Sharp drop at 22 km

Setup:

- Channel emulator
- RF isolated experiments

Throughput vs distance again



- Atheros Radios
- UDP throughput
- With MAC retries

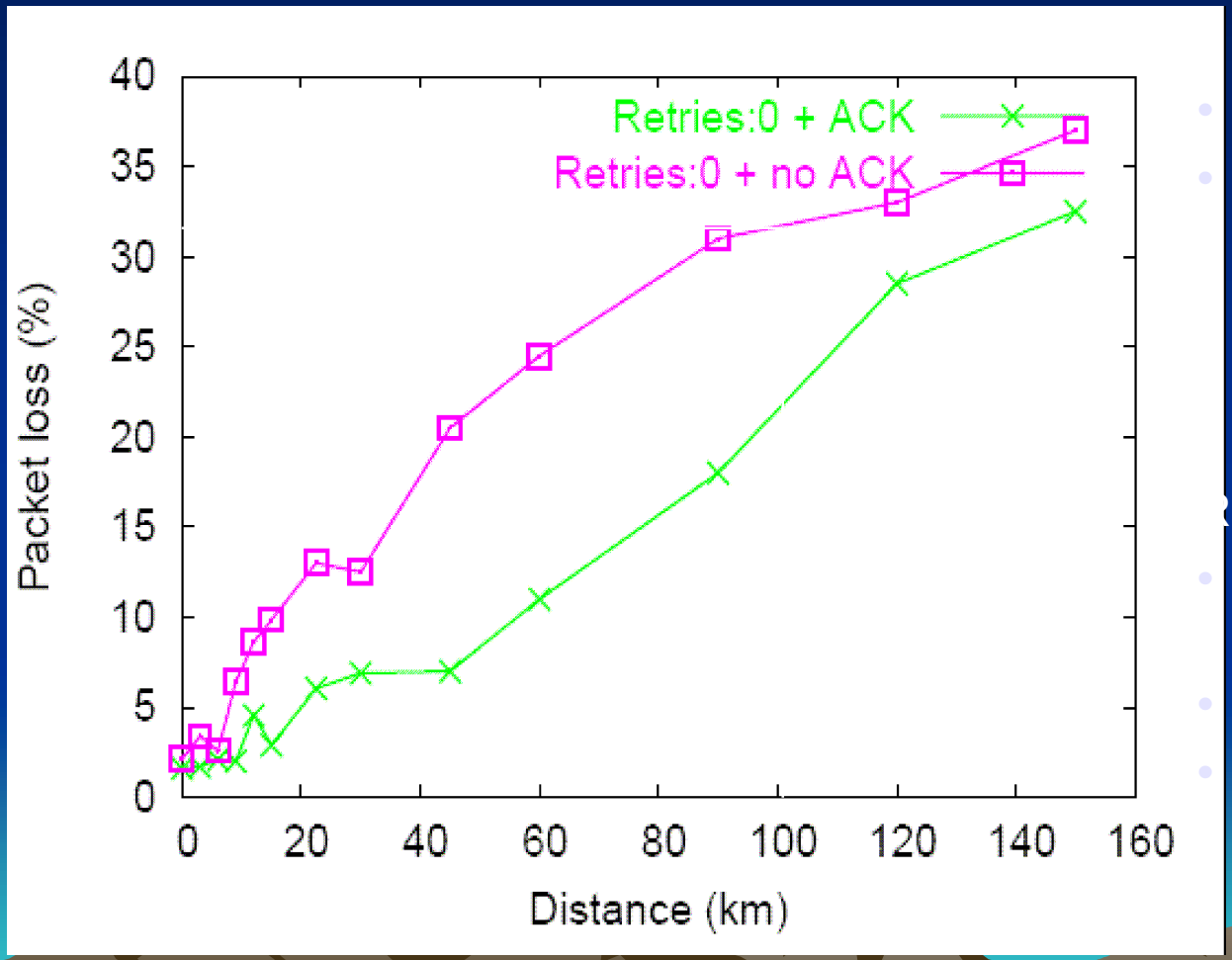
Tweak:

- Use longer ACK timeouts

Results:

- 30 % drop even before 111 km
- Sharp drop at 111 km

Bi-directional traffic vs distance



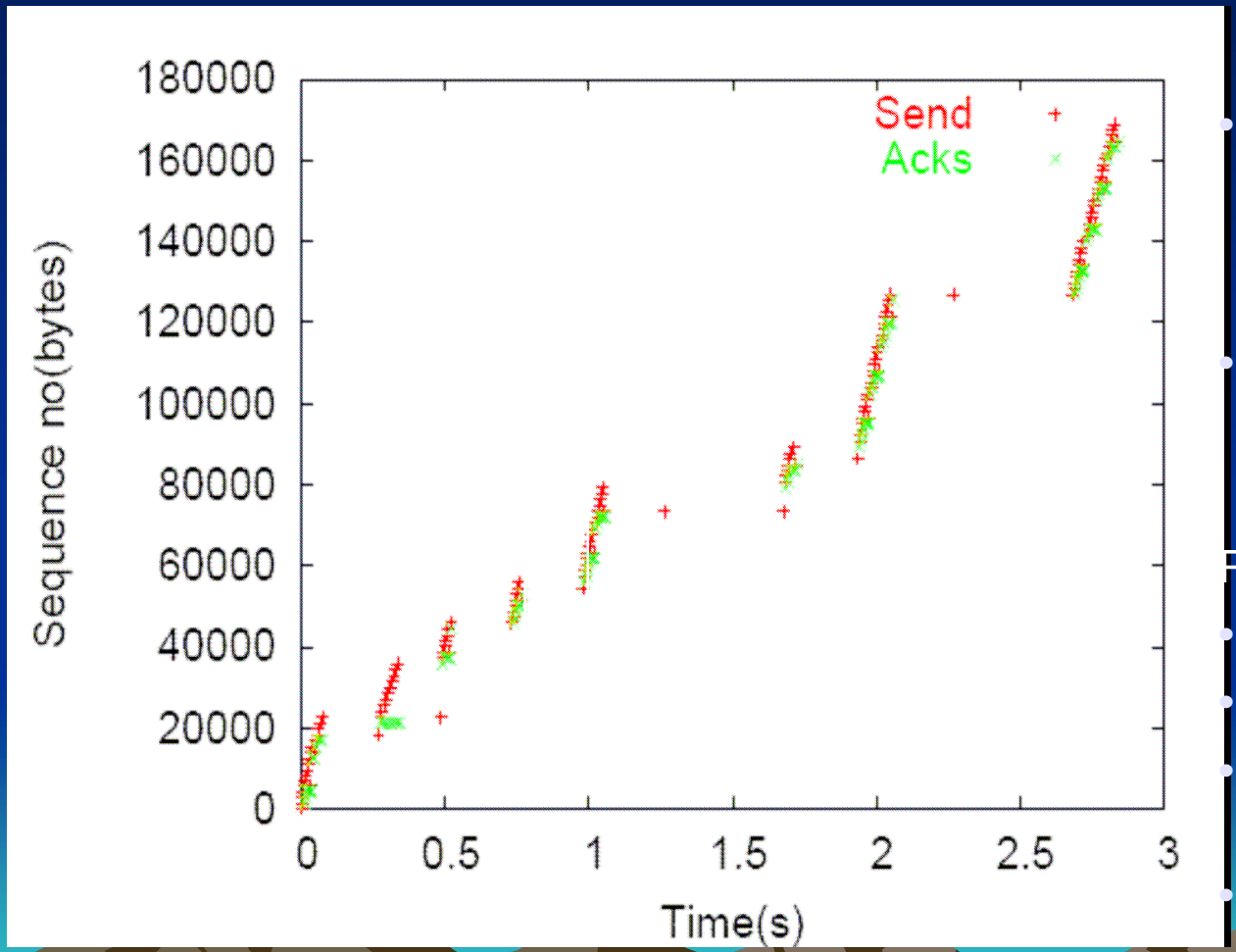
- Bi-directional UDP
- With zero MAC retries

Results:

- Collision avoidance does not work
- Too far apart
- Upto 30% at 100km

Intel Research Berkeley

TCP on Lossy link



- 20 km link over the bay from Berkeley to SF (near Sutro Tower)
- 20 % loss

Results:

- TCP : 372 kbps
- UDP : 4.5 mbps
- Frequent timeouts of 200 ms
- Lots of ACKs are also lost

But this only works if…

- End-to-end RTT is not terribly large
  - A few seconds at the very most [typ < 500ms]
  - (TCP works)
- Some path exists between endpoints
  - Routing usually finds single “best” existing route
  - (Internet Routing works)
- Retransmission enhances reliability [TCP]
  - True for low loss rates (under 2% or so)
- Packet switching is the right abstraction
  - Internet/IP makes packet switching interoperable

Non-Internet-Like Networks

- Things that move
  - Ships, gliders, AUVs, submarines, satellites…
  - Buses, trucks, bicycles, mules, zebras, etc.
- "Exotic" links
  - NASA DSN [40min max RTT to Mars]
  - UAM's [acoustics: low capacity, high error rates & latencies, intermittent, (power)]
  - Optical: free-space, underwater, etc.

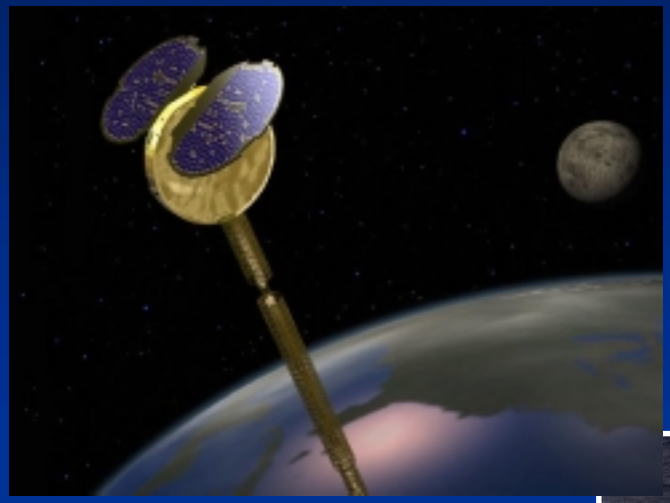
Delay-Tolerant Networking Architecture

- Goals
  - Support interoperability across ‘radically heterogeneous’ networks
  - Acceptable performance in high loss/delay/error/disconnected environments
  - Decent performance for low loss/delay/errors
- Components
  - Flexible naming scheme with *late binding*
  - Message overlay abstraction and API
  - Routing and link/contact scheduling w/CoS
  - Per-(overlay)-hop reliability and authentication

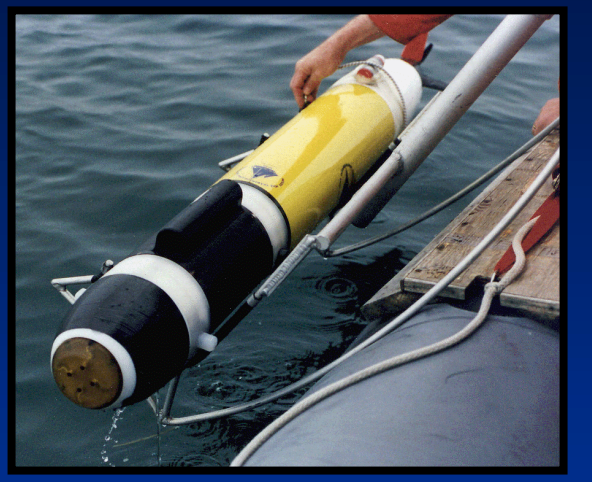
Platforms

CODAR BUOY

ORBCOMM LEO



REMUS AUV



NASA – AVIRIS

SLOCUM GLIDER

R/V ENDEAVOR

Intel Research Berkeley

http://www.dtnrg.org

http://tier.cs.berkeley.edu

Intel Research Berkeley

Oceanographic Scenario [A. Maffei/WHOI]

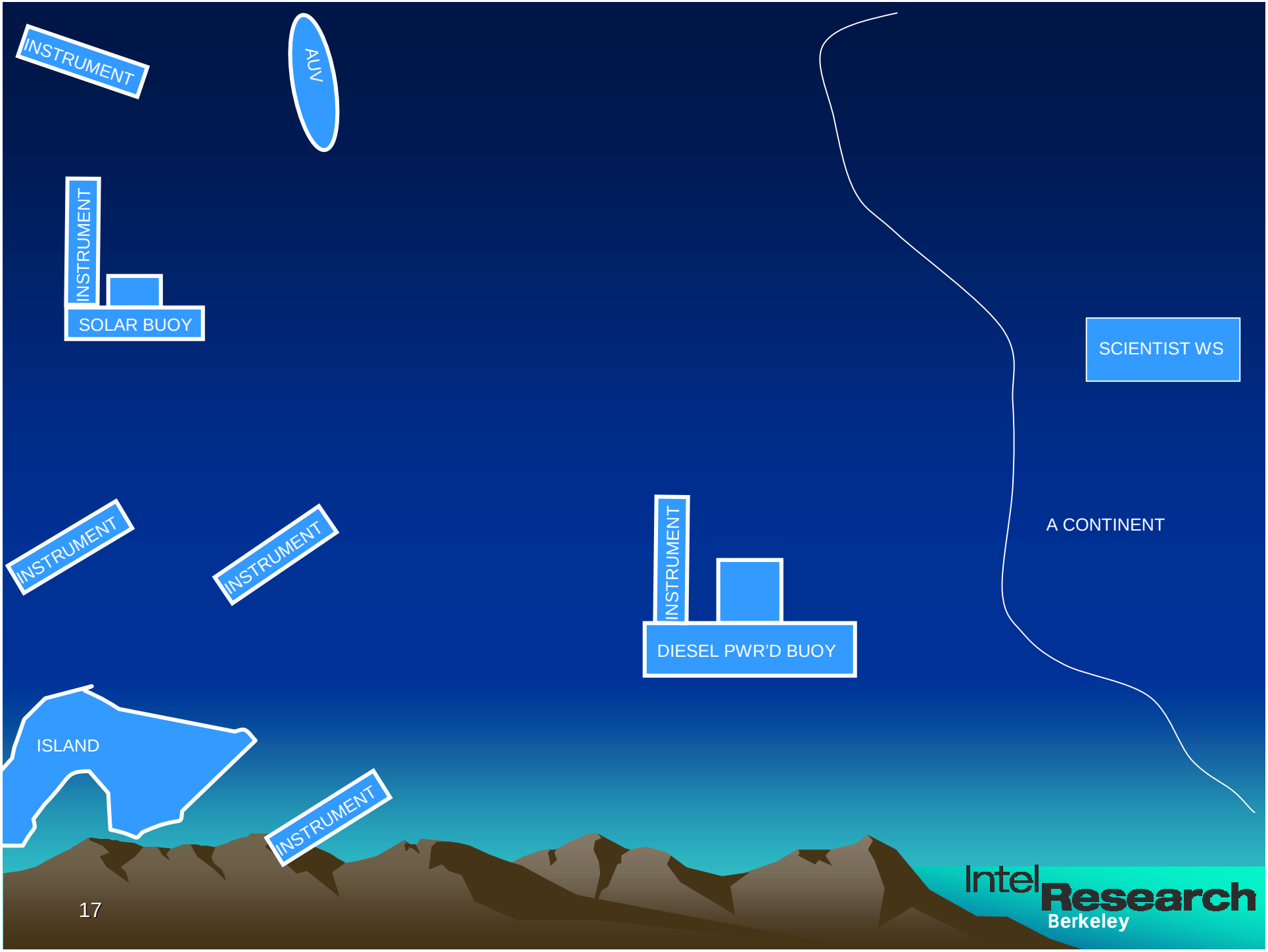
- Imagine you are an oceanographer interested in undersea earthquakes. You have deployed instrument clusters of acoustic, chemical, seismic, etc. sensors with acoustic modems attached to them.
- Autonomous Underwater Vehicles (AUVs) gather data from the instruments on a regular basis. Sometimes you make special requests to gather data from specific instruments, for example after an underwater earthquake occurs.

Scientist Interactions with Instrument Platforms

- As an oceanographer you don't concern yourself about which data communications assets are used to deliver commands to instruments or how data is delivered back to you.
- For example, a *dtn_file_copy* or *dtn_traceroute* command might be issued from your workstation command line indicating a DTN region and an oceanographic instrument name inside that region.
- The command might complete 3 weeks later when email arrives telling you that the data has arrived or displaying a report showing the communications path and sample round-trip latency to the instrument(s) you are interested in.

Potential Oceanographic DTN Regions

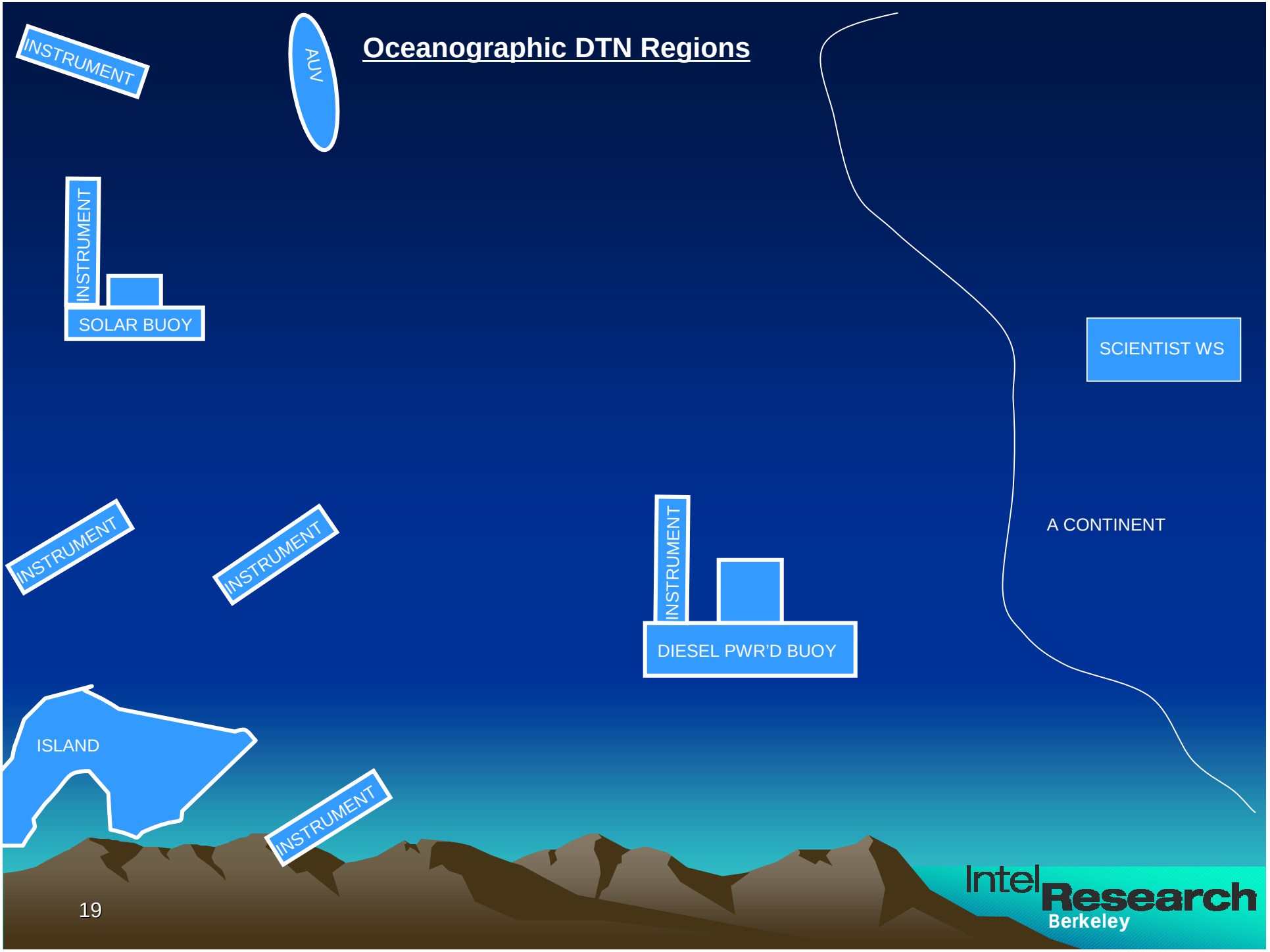
- First, the platforms of interest



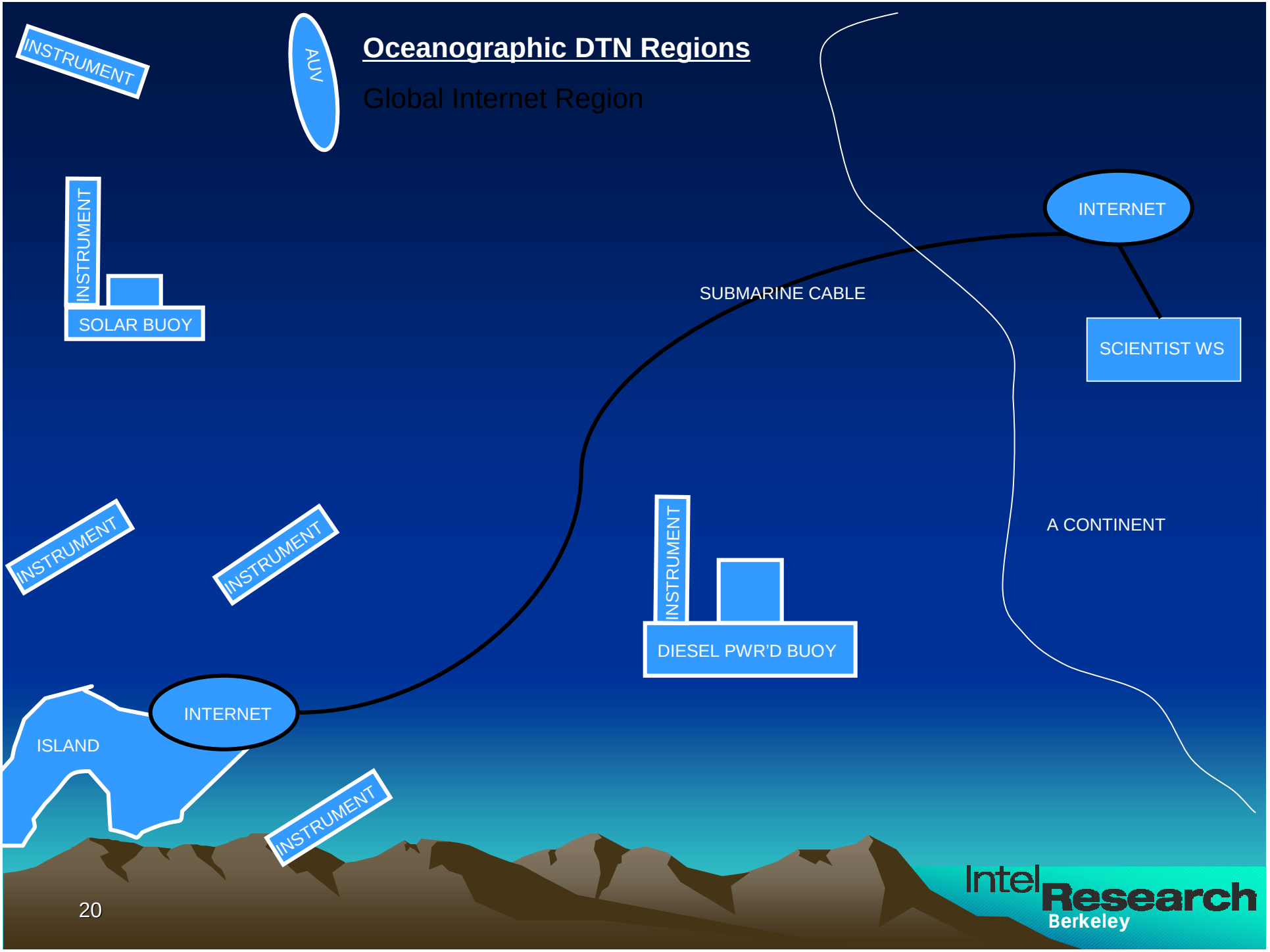
Intel Research Berkeley

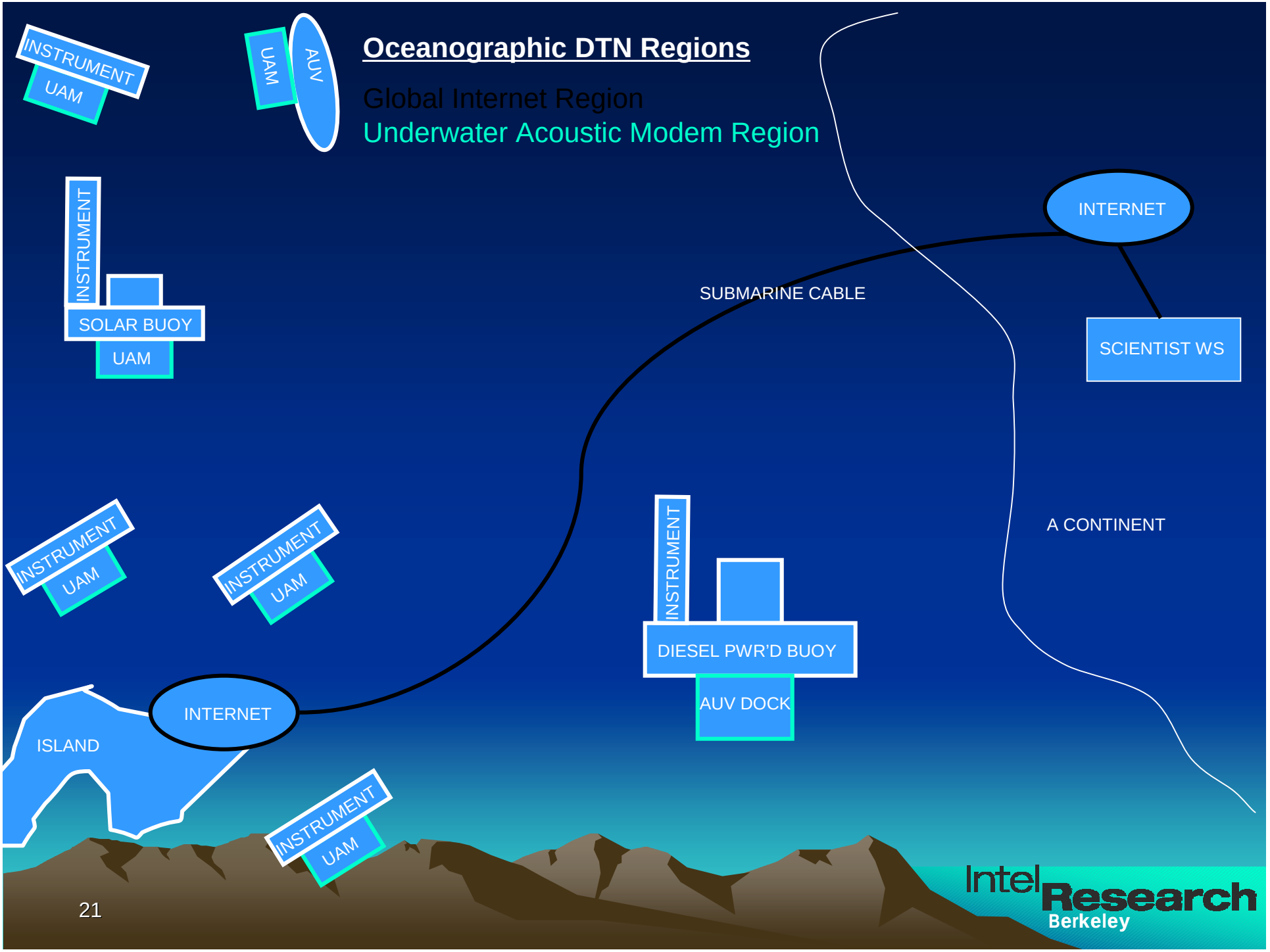
Potential Oceanographic DTN Regions

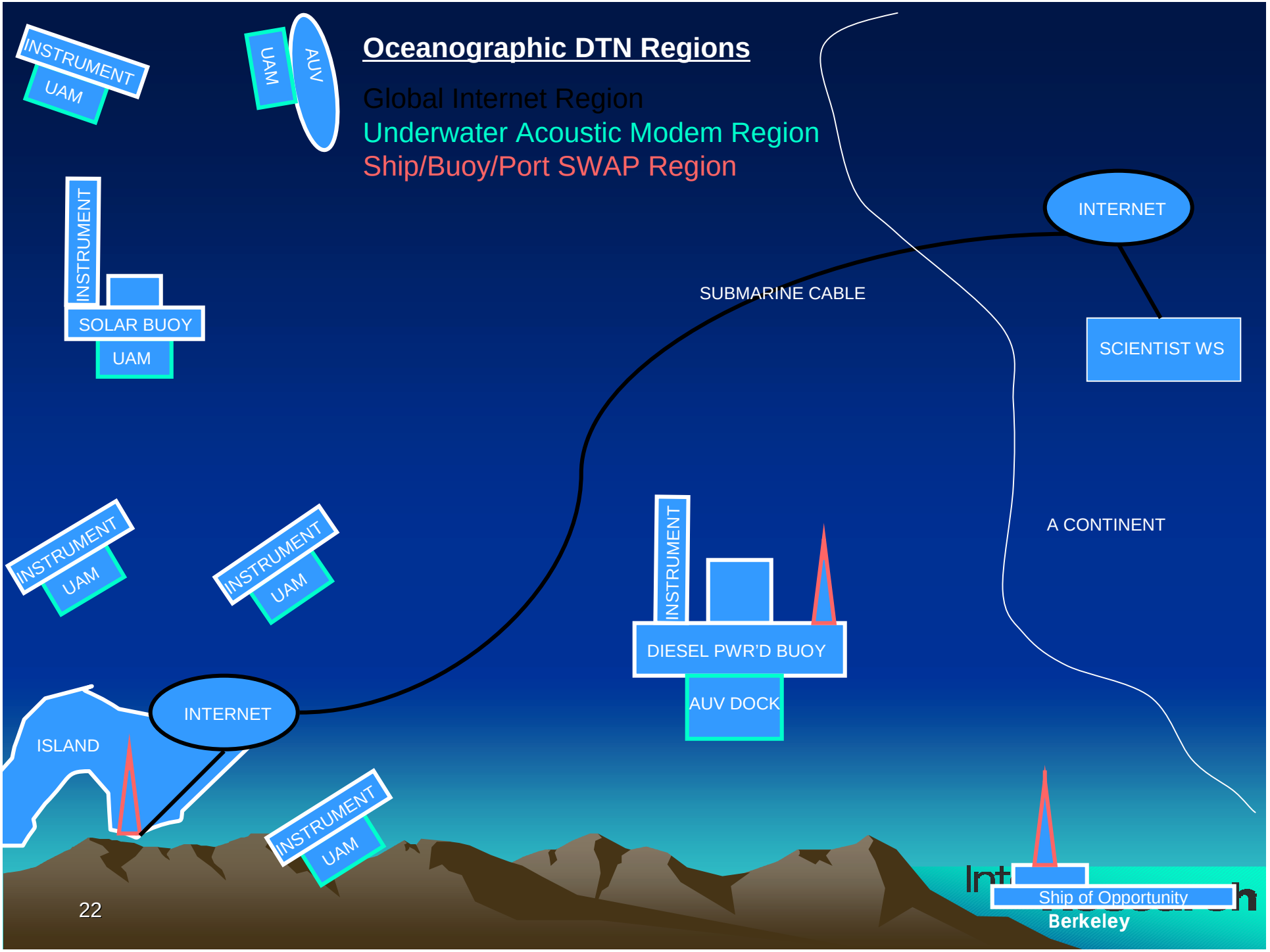
- First, the platforms of interest
- Next, some Oceanographic DTN regions that could be developed (like address families).

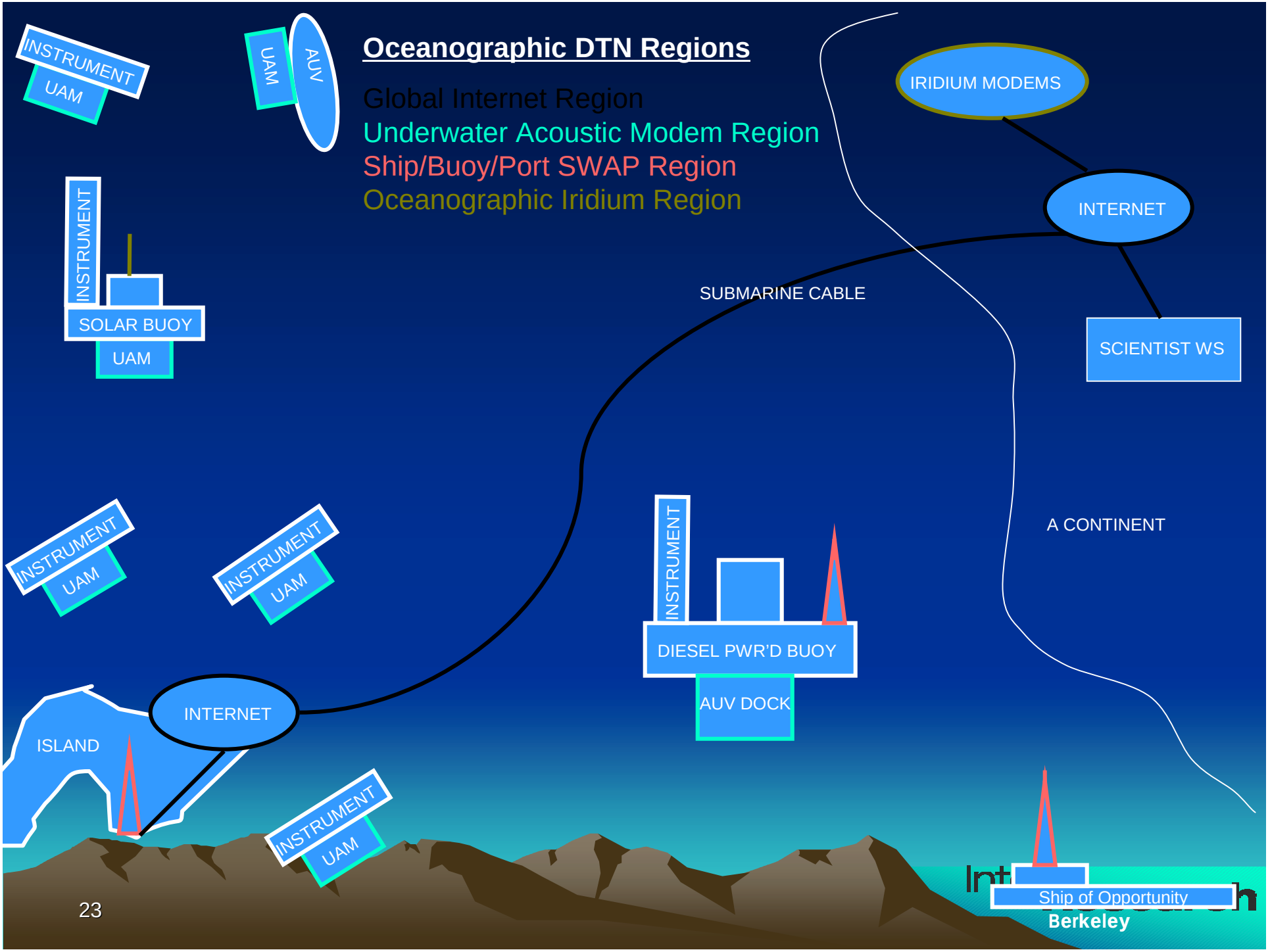


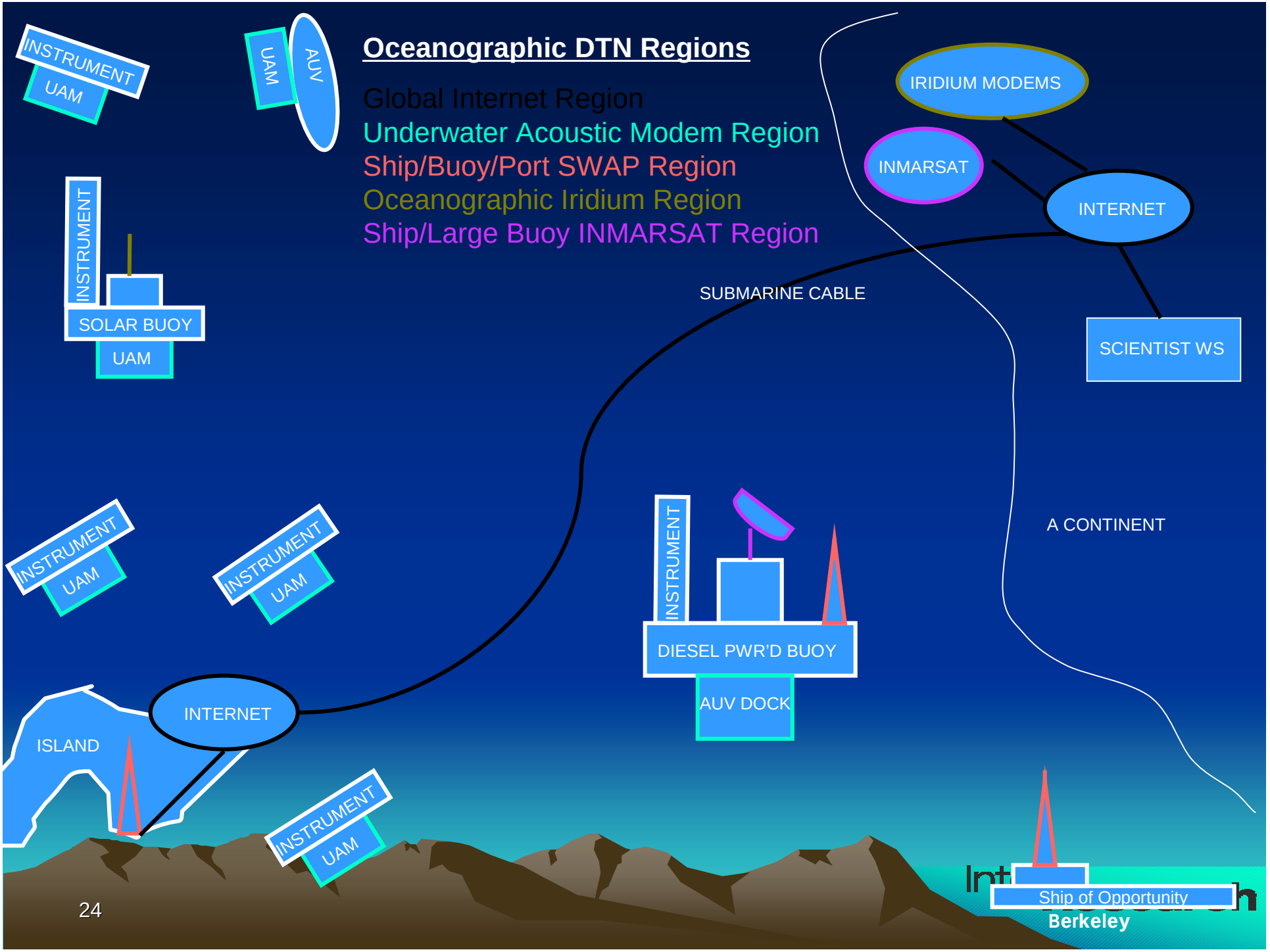
Intel Research Berkeley


Oceanographic DTN Regions
Global Internet Region
INSTRUMENT
AUV
INSTRUMENT
SOLAR BUOY
INTERNET
SUBMARINE CABLE
SCIENTIST WS
INSTRUMENT
INSTRUMENT
INSTRUMENT
DIESEL PWR'D BUOY
A CONTINENT
INTERNET
ISLAND
INSTRUMENT
Intel Research
Berkeley


Oceanographic DTN Regions
Global Internet Region
Underwater Acoustic Modem Region
INSTRUMENT
UAM
UAM
AUV
INSTRUMENT
SOLAR BUOY
UAM
INTERNET
SUBMARINE CABLE
SCIENTIST WS
INSTRUMENT
UAM
INSTRUMENT
UAM
INSTRUMENT
DIESEL PWR'D BUOY
AUV DOCK
A CONTINENT
INTERNET
ISLAND
INSTRUMENT
UAM
Intel Research
Berkeley

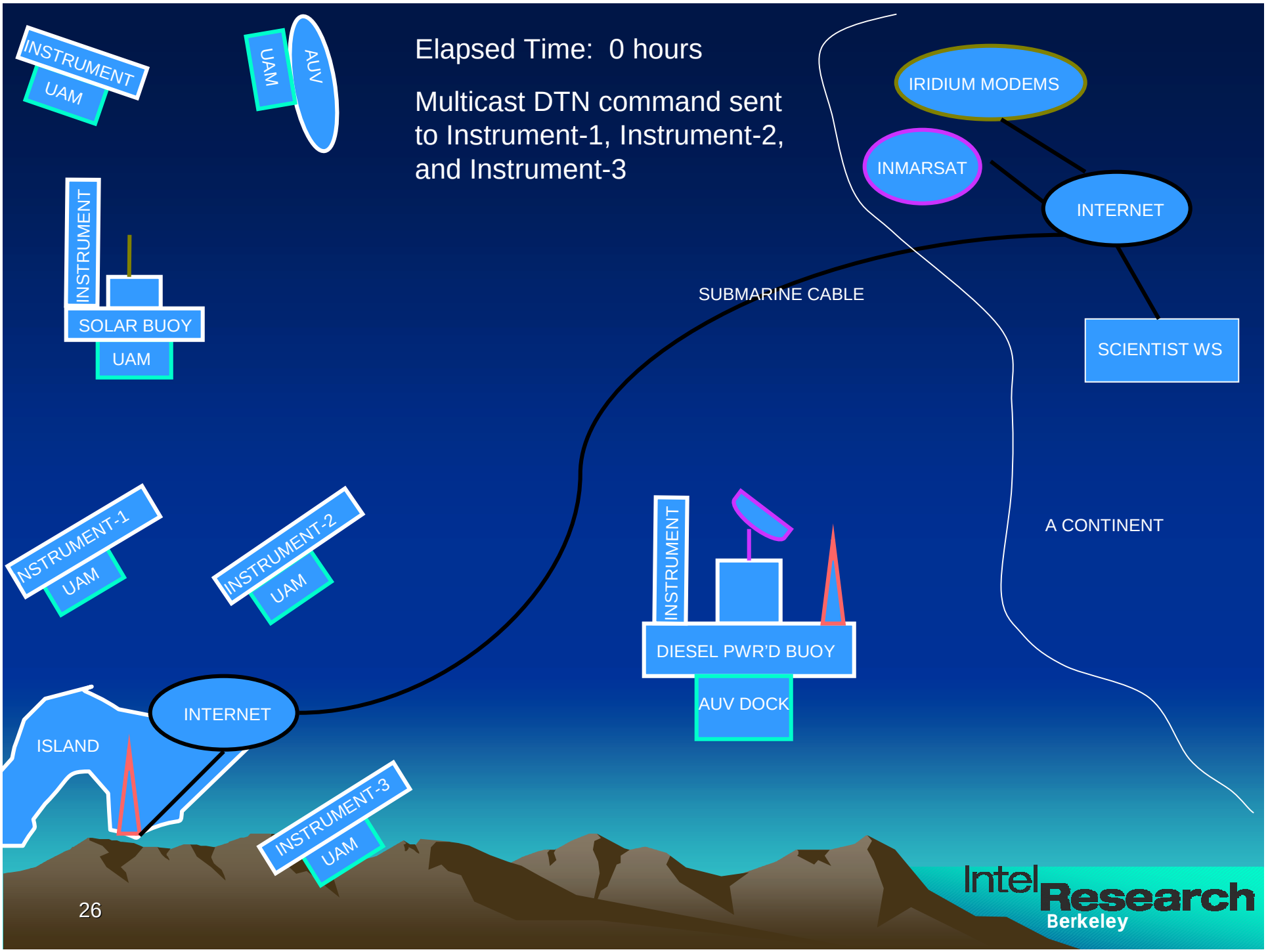

Oceanographic DTN Regions
Global Internet Region
Underwater Acoustic Modem Region
Ship/Buoy/Port SWAP Region
INSTRUMENT
UAM
UAM
AUV
INSTRUMENT
SOLAR BUOY
UAM
INTERNET
SUBMARINE CABLE
SCIENTIST WS
INSTRUMENT
UAM
INSTRUMENT
UAM
INSTRUMENT
DIESEL PWR'D BUOY
AUV DOCK
A CONTINENT
INTERNET
ISLAND
INSTRUMENT
UAM
Ship of Opportunity
Berkeley


Oceanographic DTN Regions
Global Internet Region
Underwater Acoustic Modem Region
Ship/Buoy/Port SWAP Region
Oceanographic Iridium Region
INSTRUMENT
UAM
UAM
AUV
IRIDIUM MODEMS
INTERNET
SUBMARINE CABLE
SCIENTIST WS
INSTRUMENT
SOLAR BUOY
UAM
INSTRUMENT
UAM
INSTRUMENT
UAM
INSTRUMENT
DIESEL PWR'D BUOY
AUV DOCK
A CONTINENT
INTERNET
ISLAND
INSTRUMENT
UAM
Ship of Opportunity
Berkeley

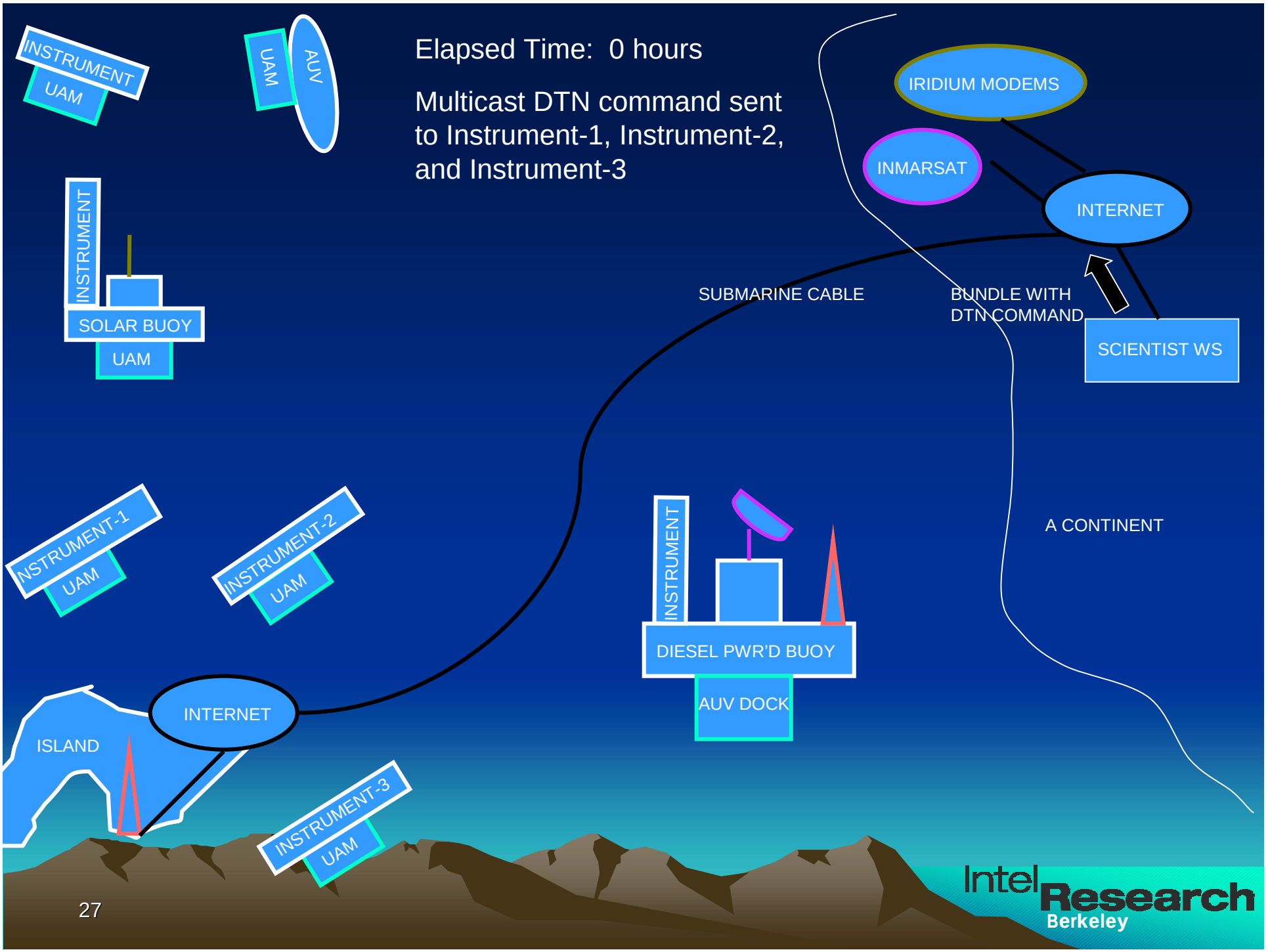

Oceanographic DTN Regions
Global Internet Region
Underwater Acoustic Modem Region
Ship/Buoy/Port SWAP Region
Oceanographic Iridium Region
Ship/Large Buoy INMARSAT Region
INSTRUMENT
UAM
UAM
AUV
IRIDIUM MODEMS
INMARSAT
INTERNET
SUBMARINE CABLE
SCIENTIST WS
INSTRUMENT
SOLAR BUOY
UAM
A CONTINENT
INSTRUMENT
UAM
INSTRUMENT
UAM
INSTRUMENT
DIESEL PWR'D BUOY
AUV DOCK
INTERNET
ISLAND
INSTRUMENT
UAM
Ship of Opportunity
Berkeley

Potential Oceanographic DTN Regions

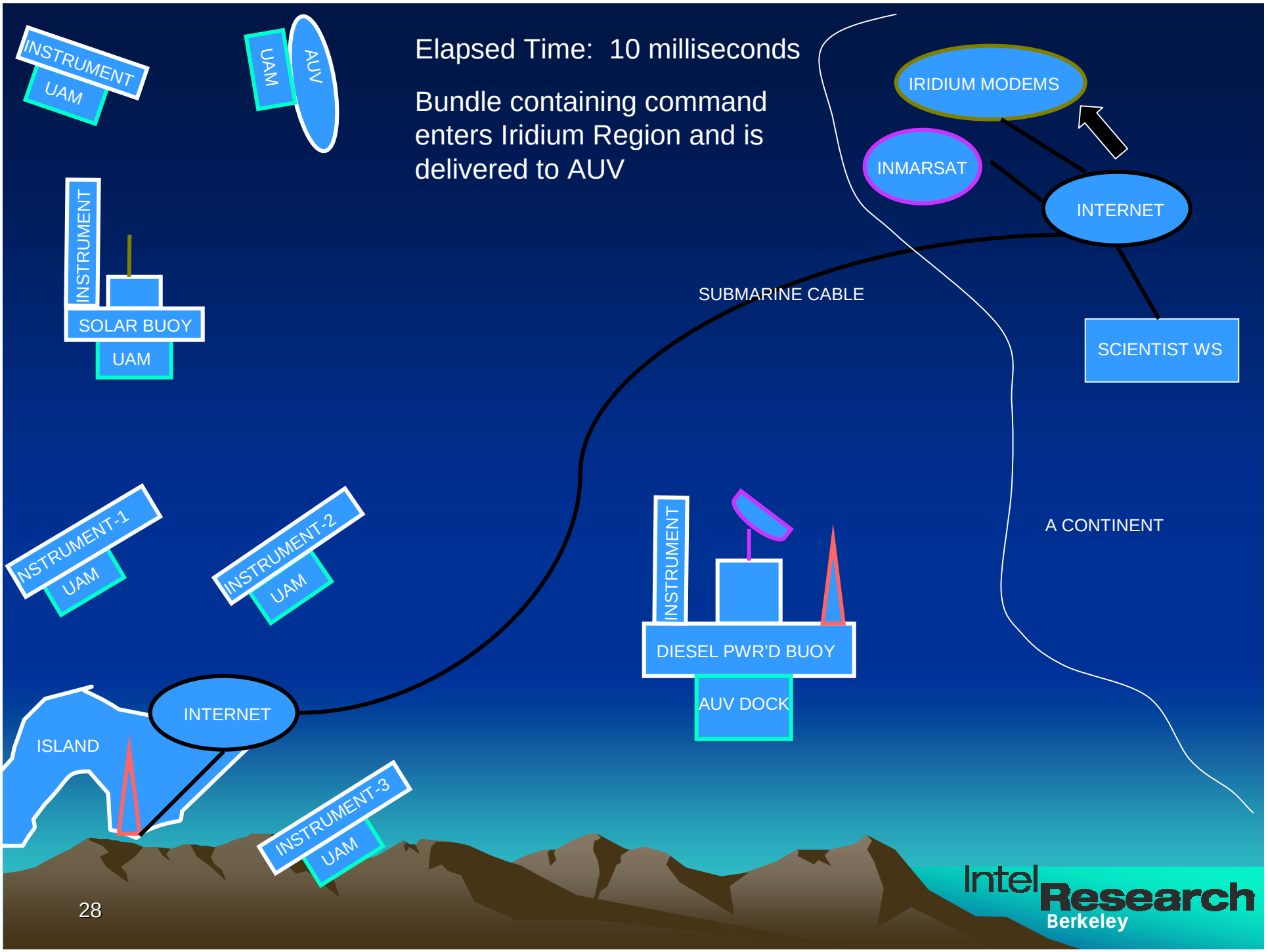
- First, the platforms of interest
- Next, some Oceanographic DTN regions that could be developed.
- A short animation



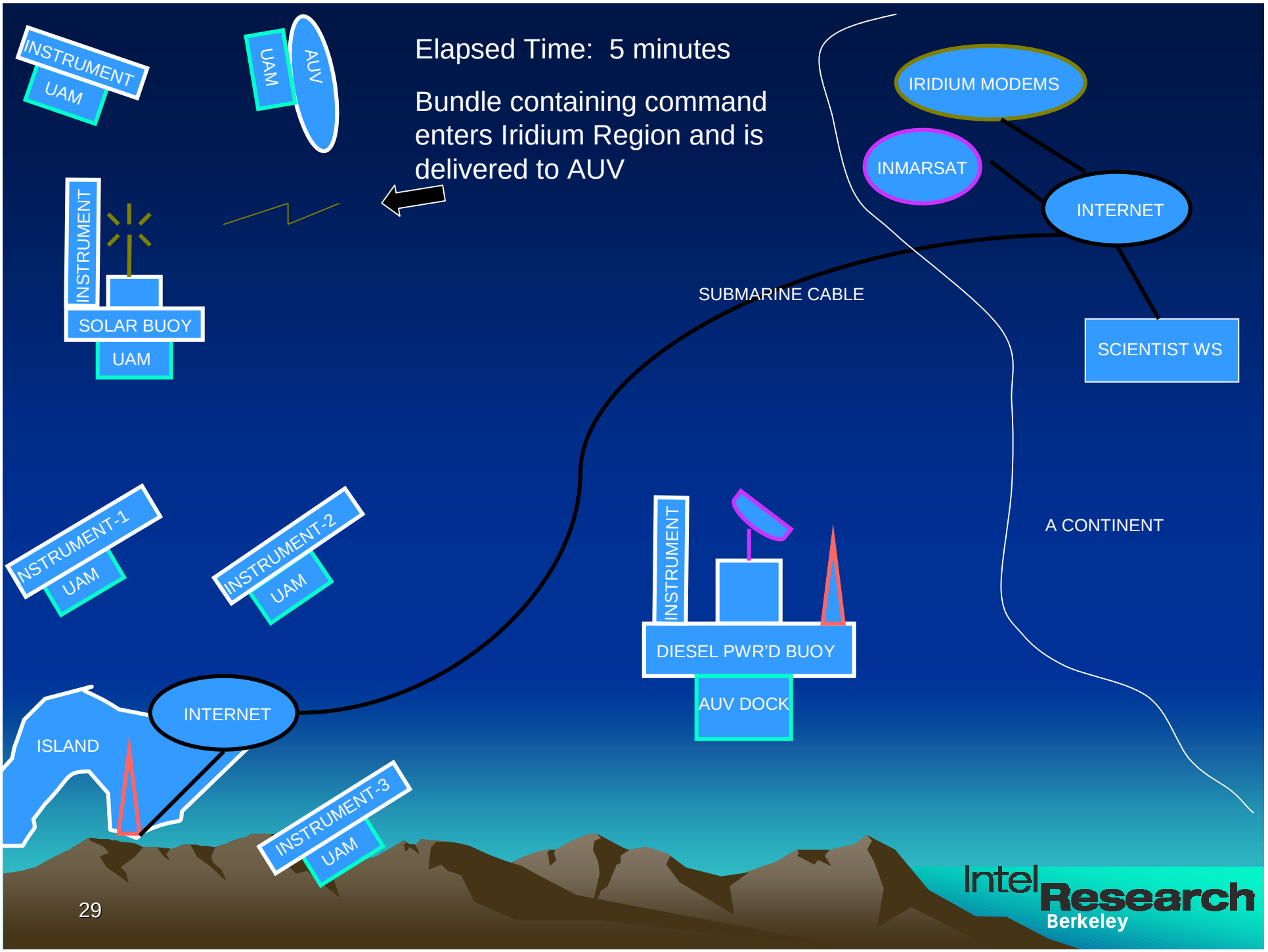
Intel Research Berkeley



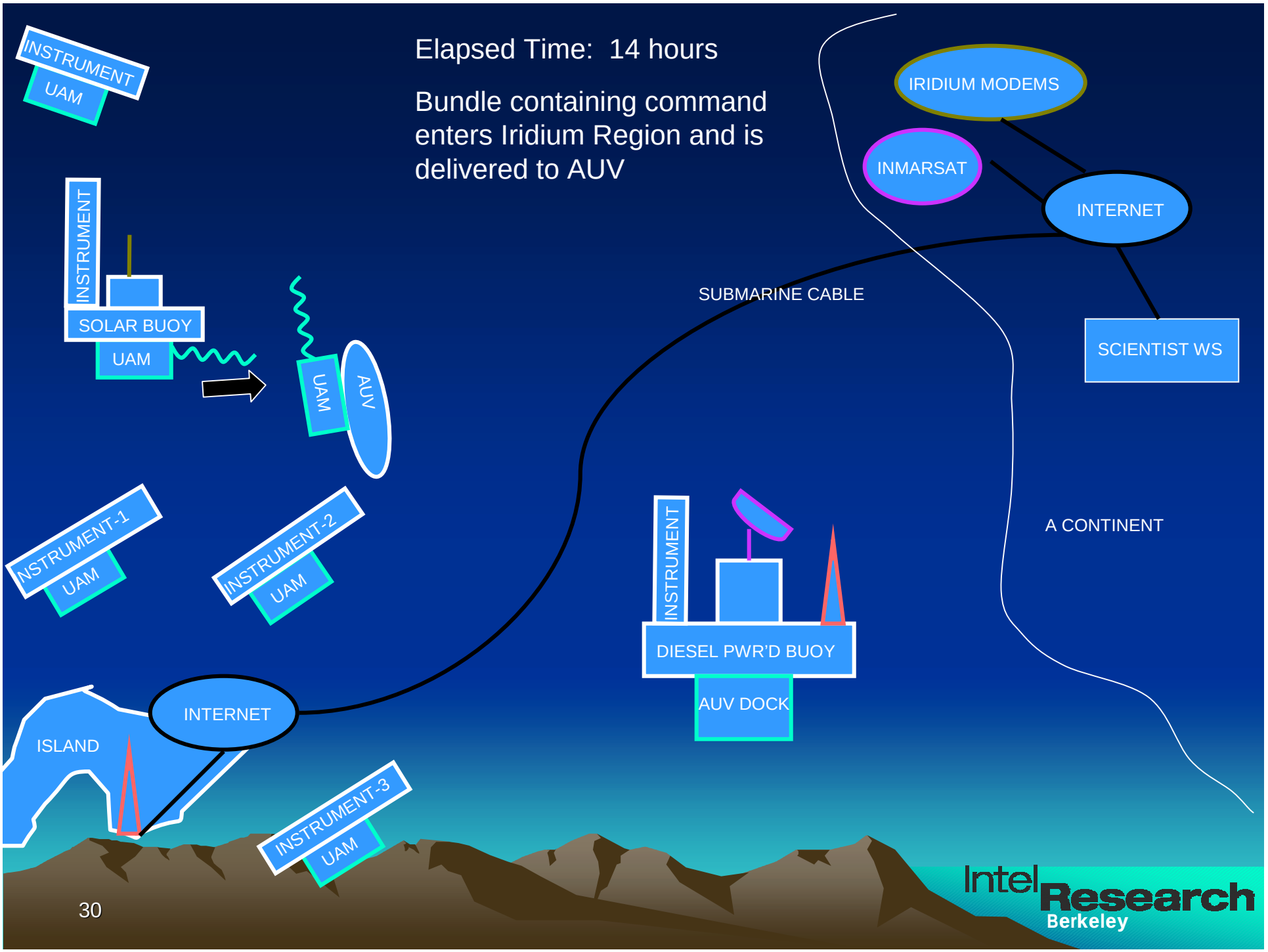
Intel Research Berkeley



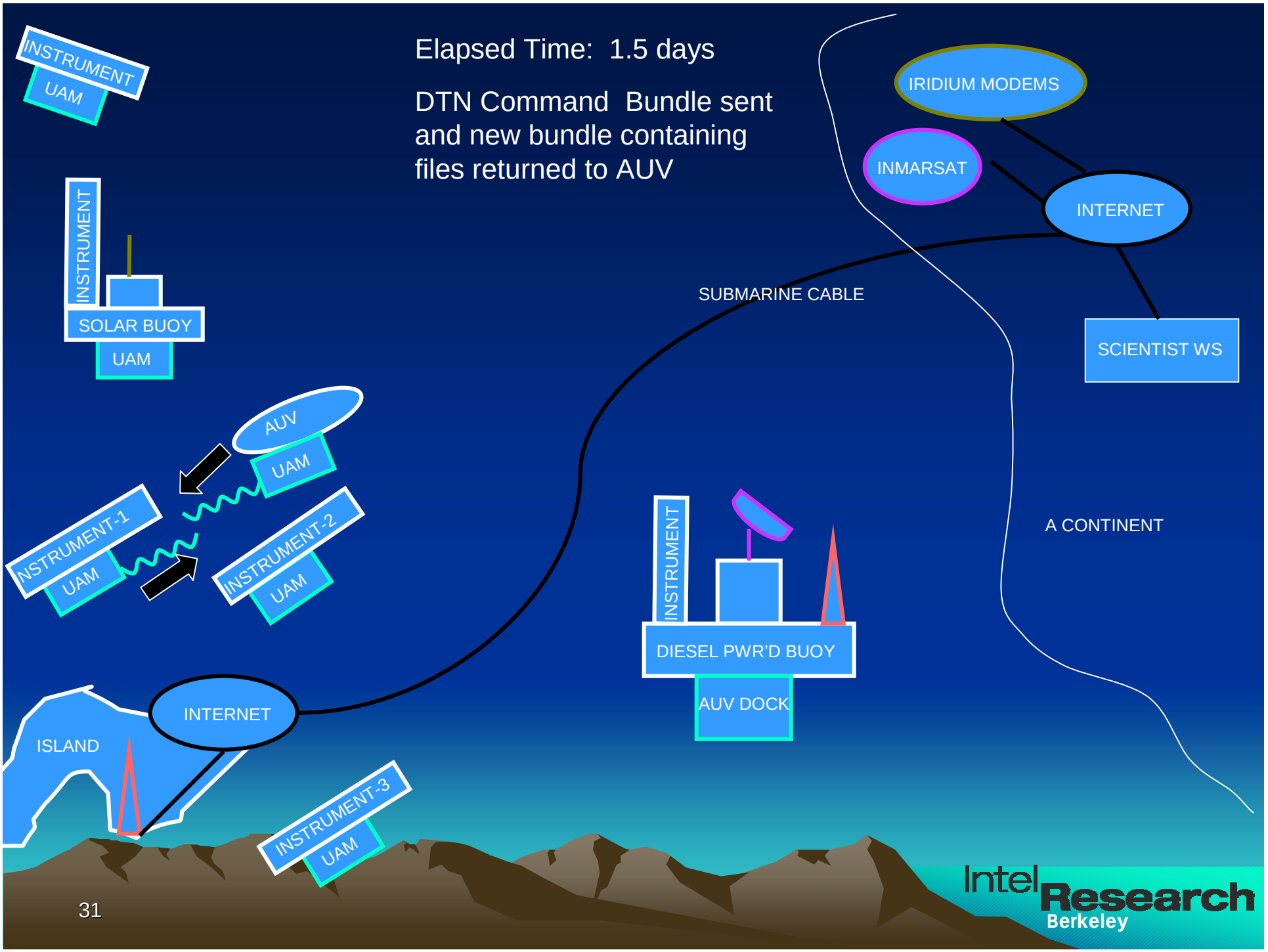
Intel Research Berkeley



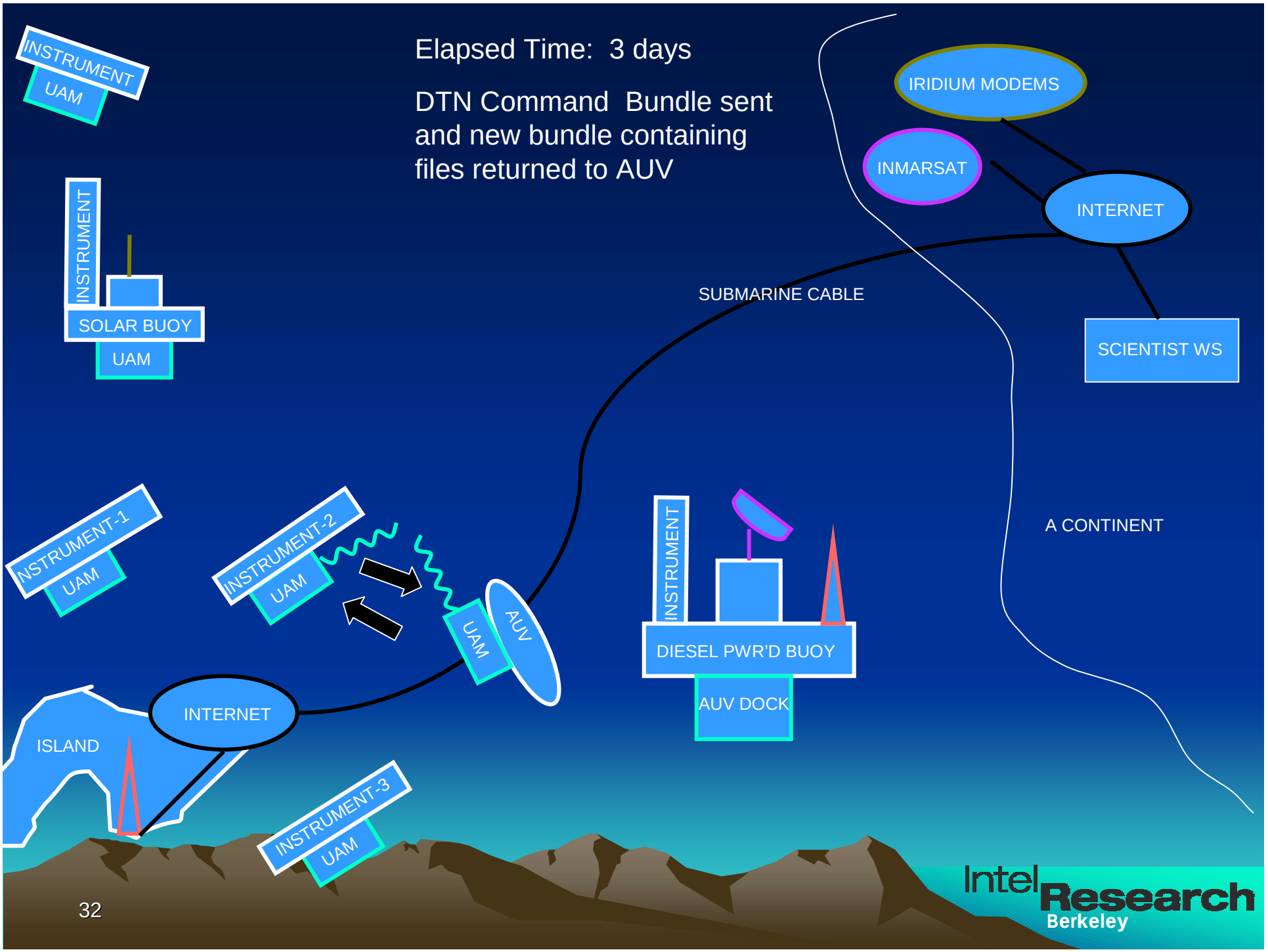
Intel Research Berkeley



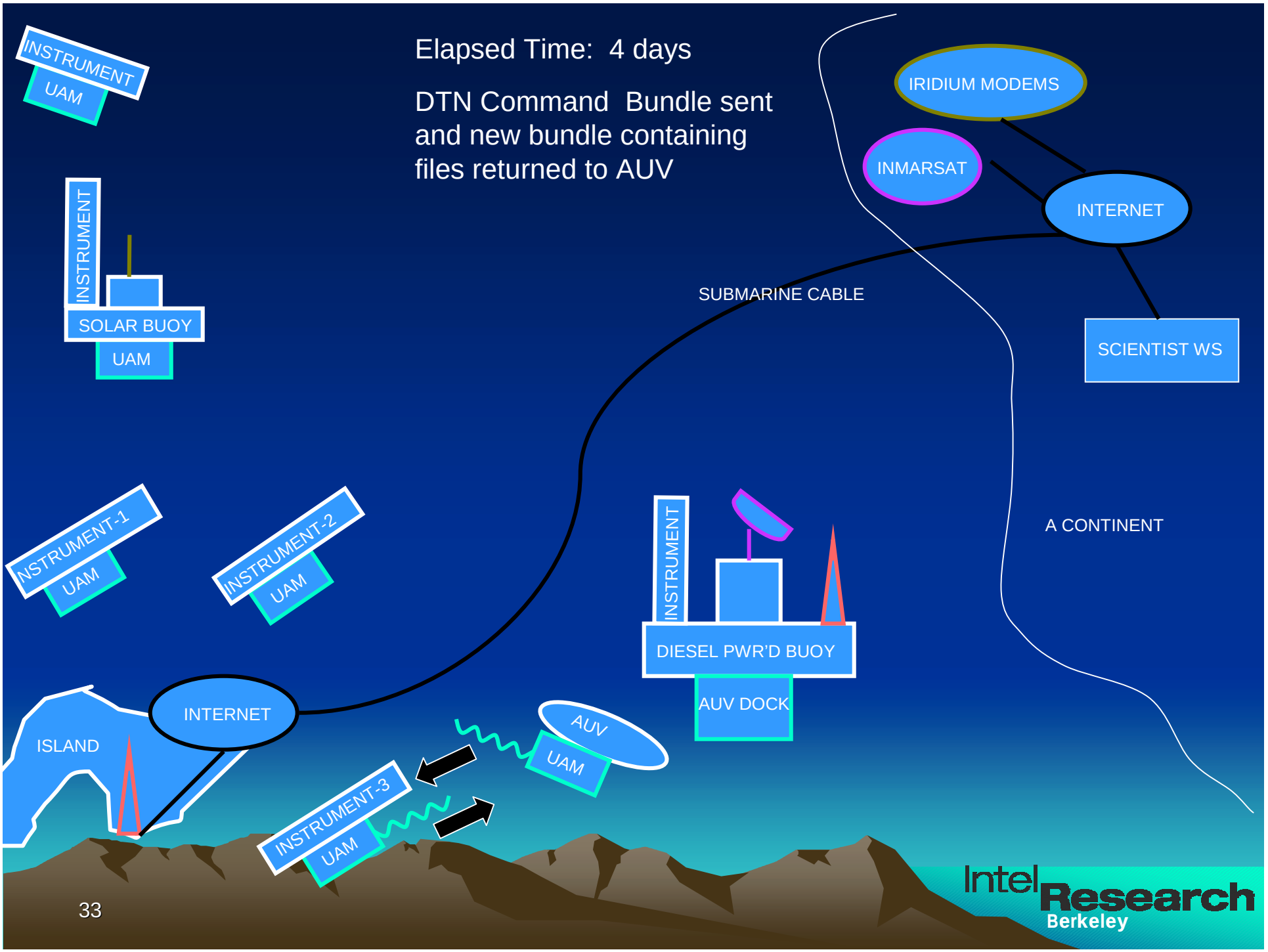
Intel Research Berkeley



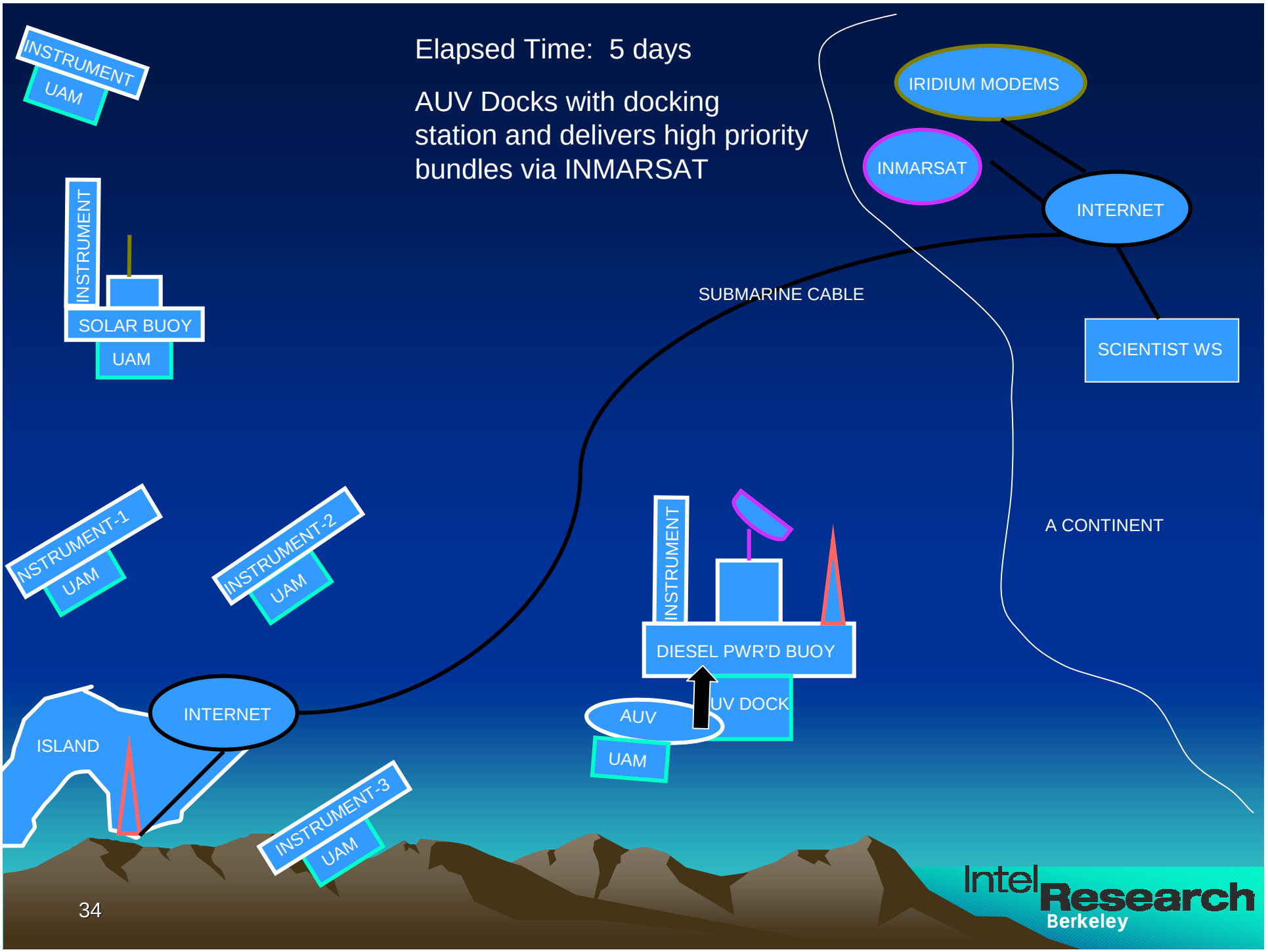
Intel Research Berkeley



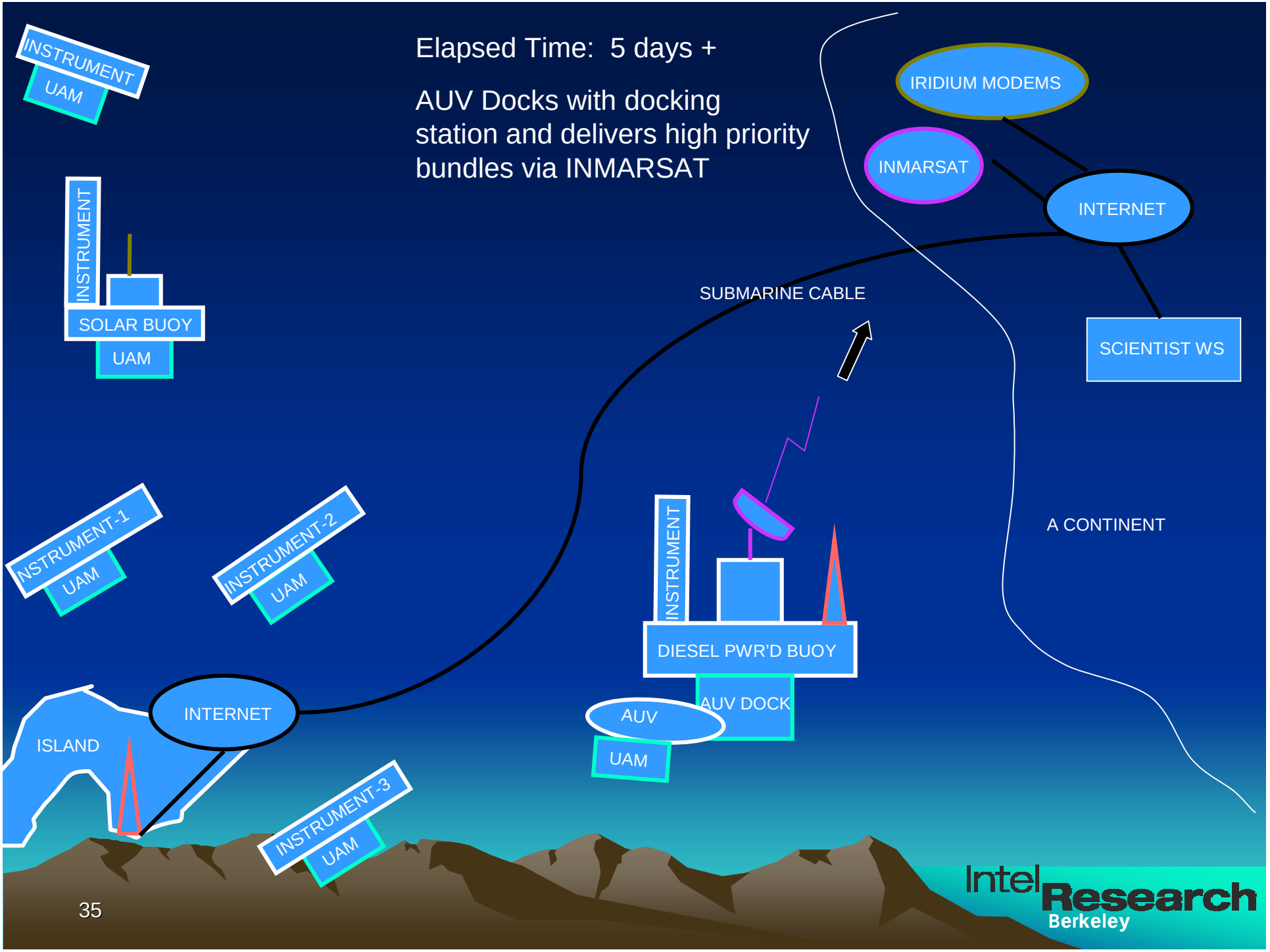
Intel Research Berkeley



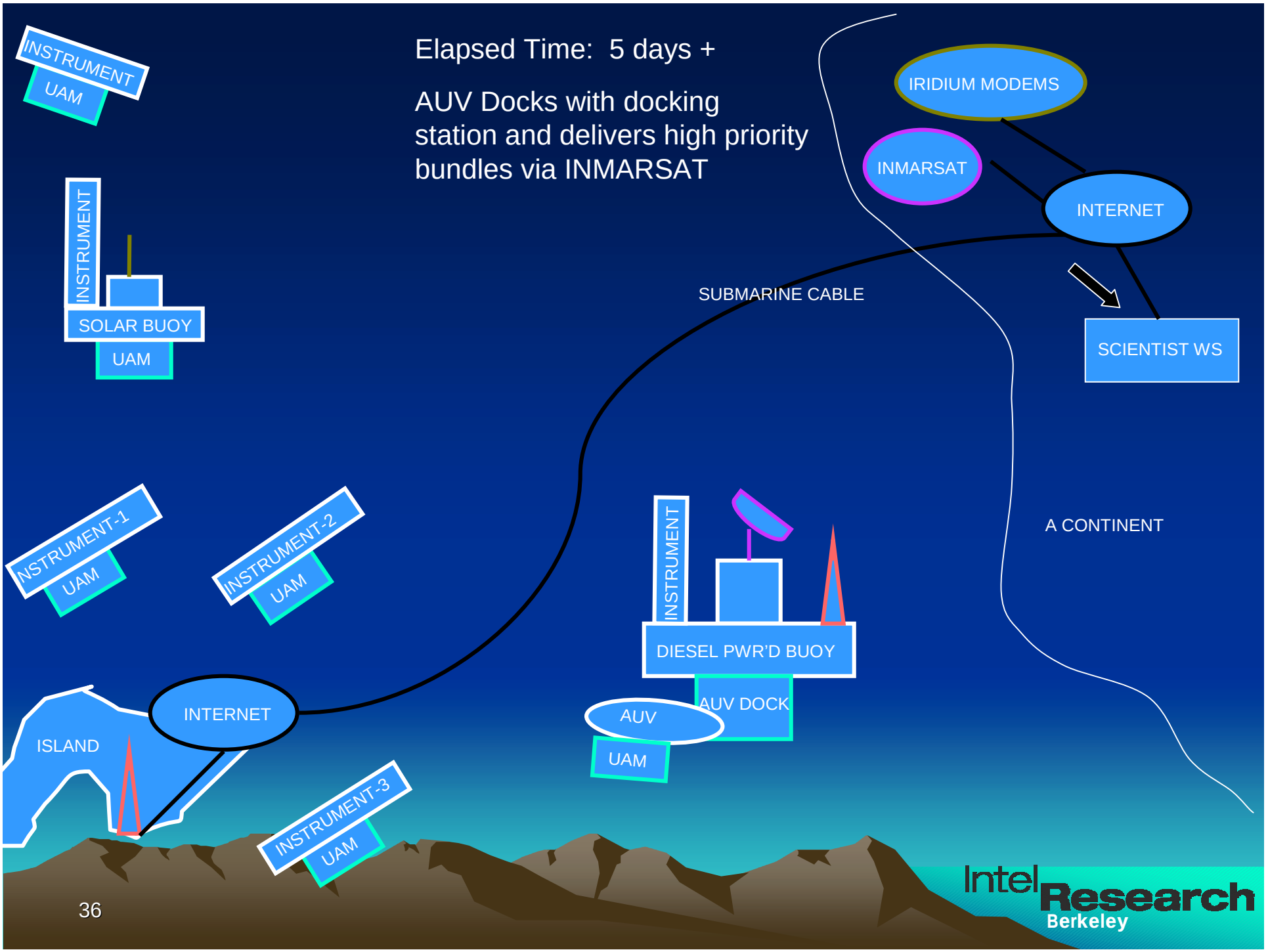
Intel Research Berkeley



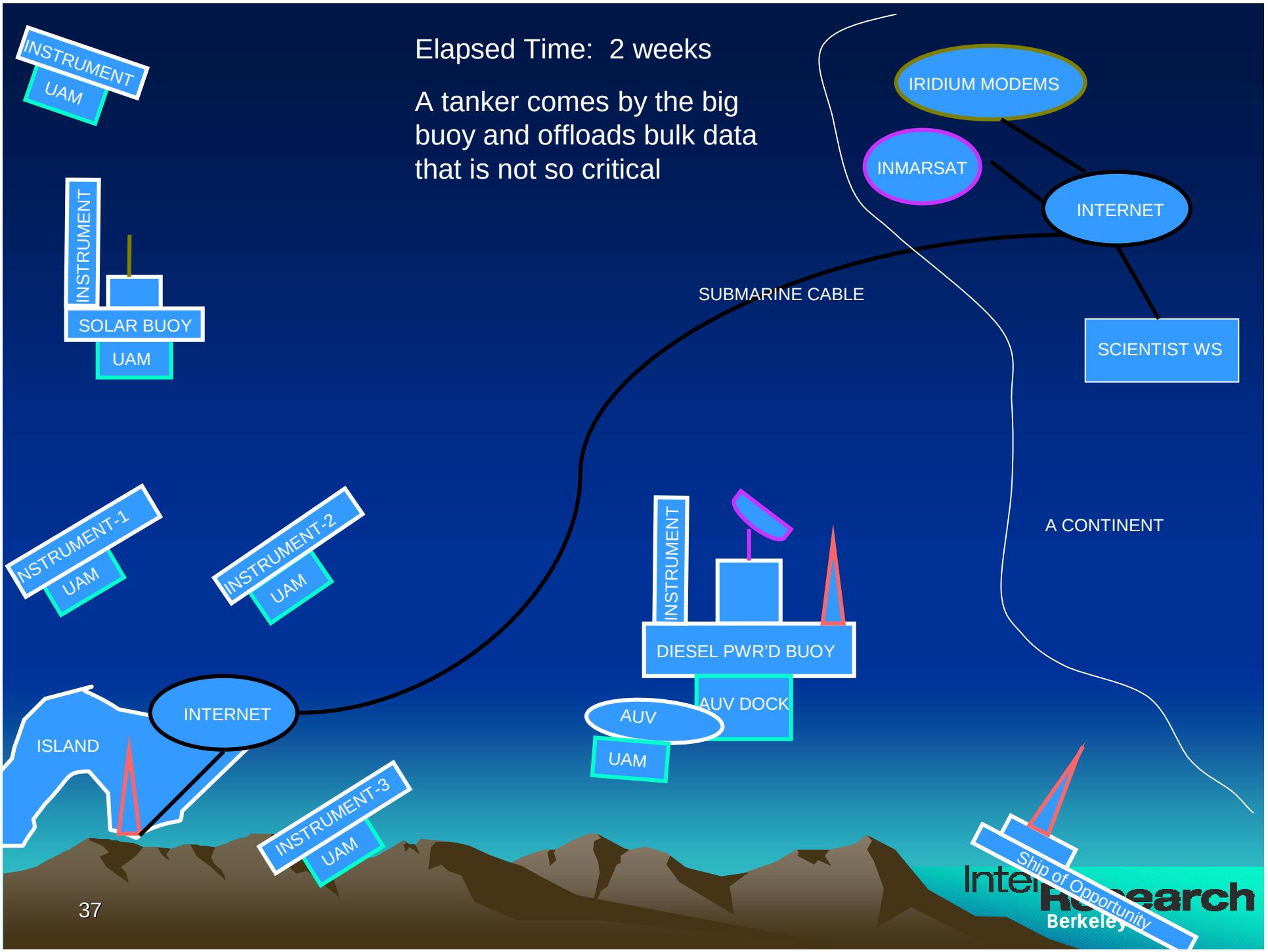
Intel Research Berkeley

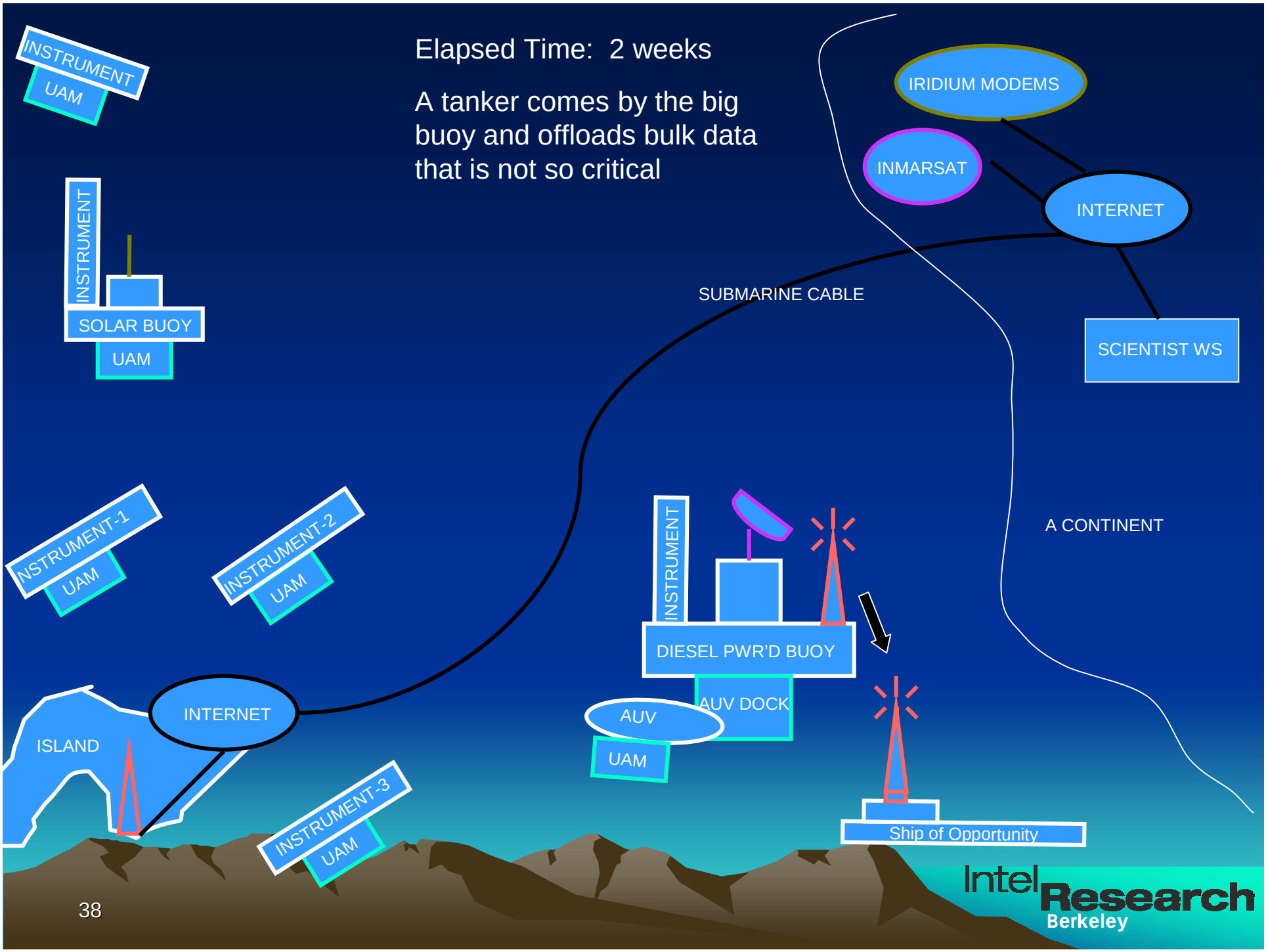


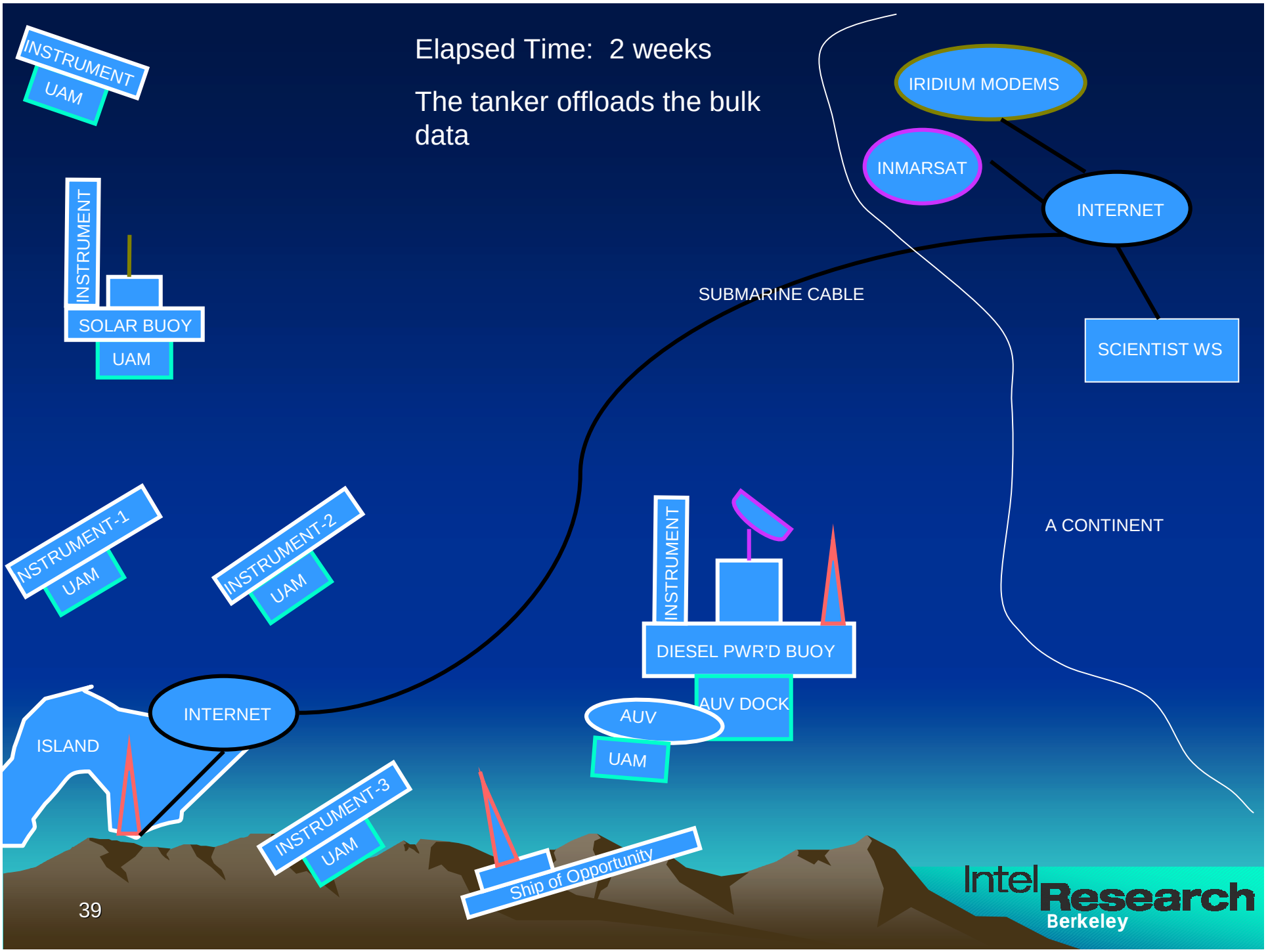
Intel Research Berkeley



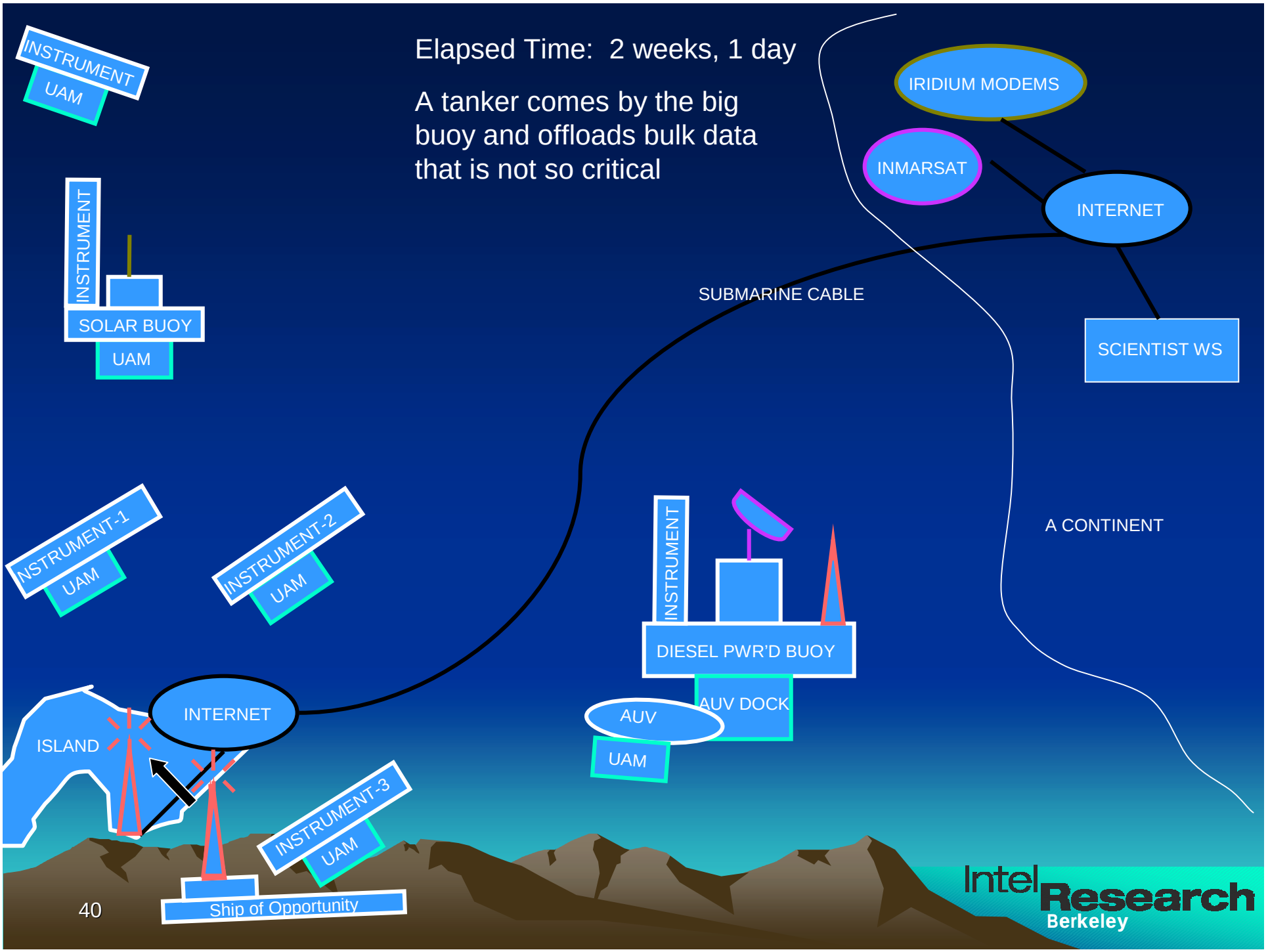
Intel Research Berkeley


Elapsed Time: 2 weeks
A tanker comes by the big buoy and offloads bulk data that is not so critical
INSTRUMENT
UAM
IRIDIUM MODEMS
INMARSAT
INTERNET
INSTRUMENT
SOLAR BUOY
UAM
SUBMARINE CABLE
SCIENTIST WS
A CONTINENT
INSTRUMENT-1
UAM
INSTRUMENT-2
UAM
INSTRUMENT
DIESEL PWR'D BUOY
AUV DOCK
AUV
UAM
INTERNET
ISLAND
INSTRUMENT-3
UAM
Ship of Opportunity
Berkeley

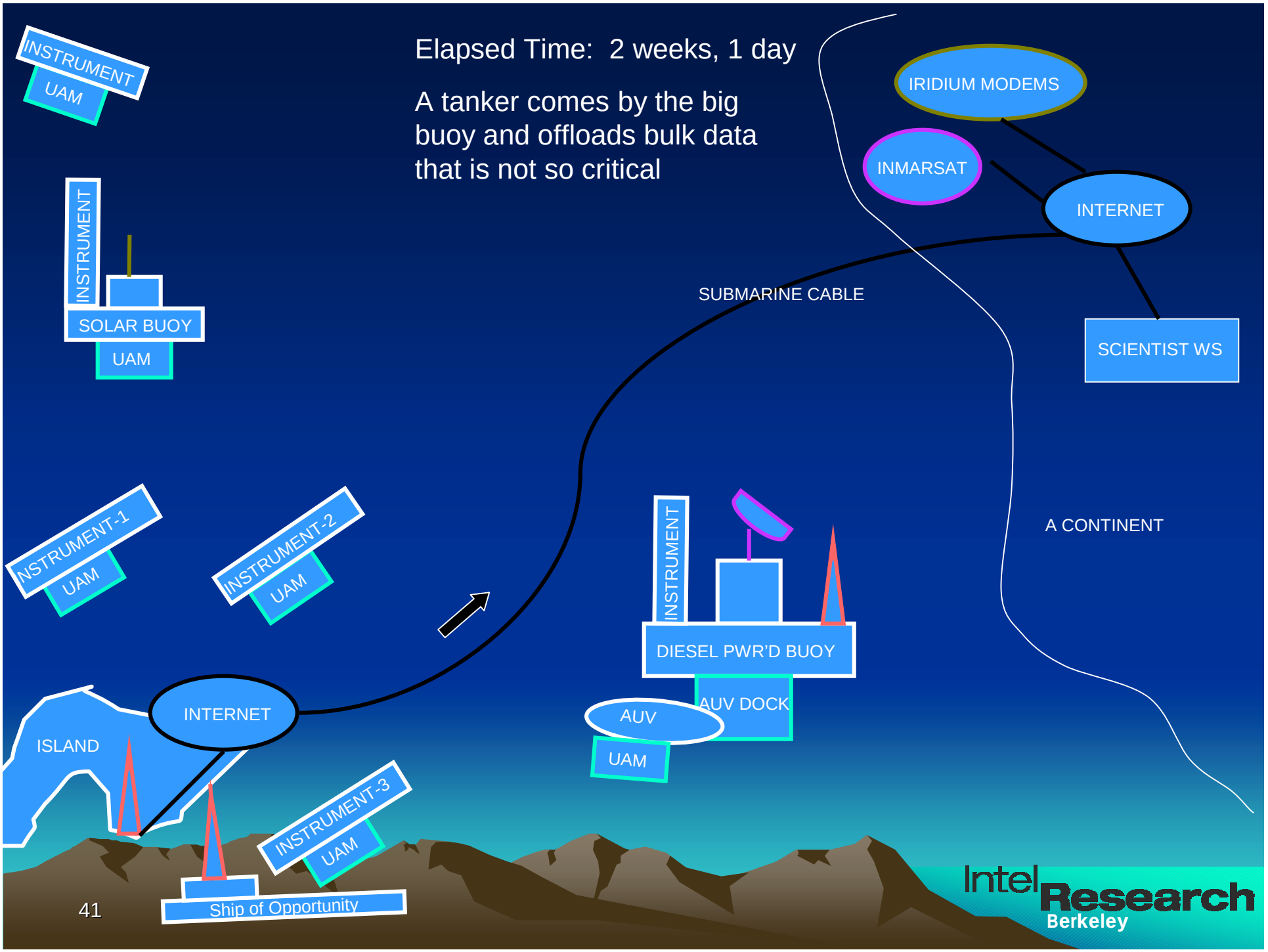

Elapsed Time: 2 weeks
A tanker comes by the big buoy and offloads bulk data that is not so critical
INSTRUMENT
UAM
IRIDIUM MODEMS
INMARSAT
INTERNET
INSTRUMENT
SOLAR BUOY
UAM
SUBMARINE CABLE
SCIENTIST WS
A CONTINENT
INSTRUMENT-1
UAM
INSTRUMENT-2
UAM
INSTRUMENT
DIESEL PWR'D BUOY
AUV DOCK
AUV
UAM
INTERNET
ISLAND
INSTRUMENT-3
UAM
Ship of Opportunity
Intel Research
Berkeley



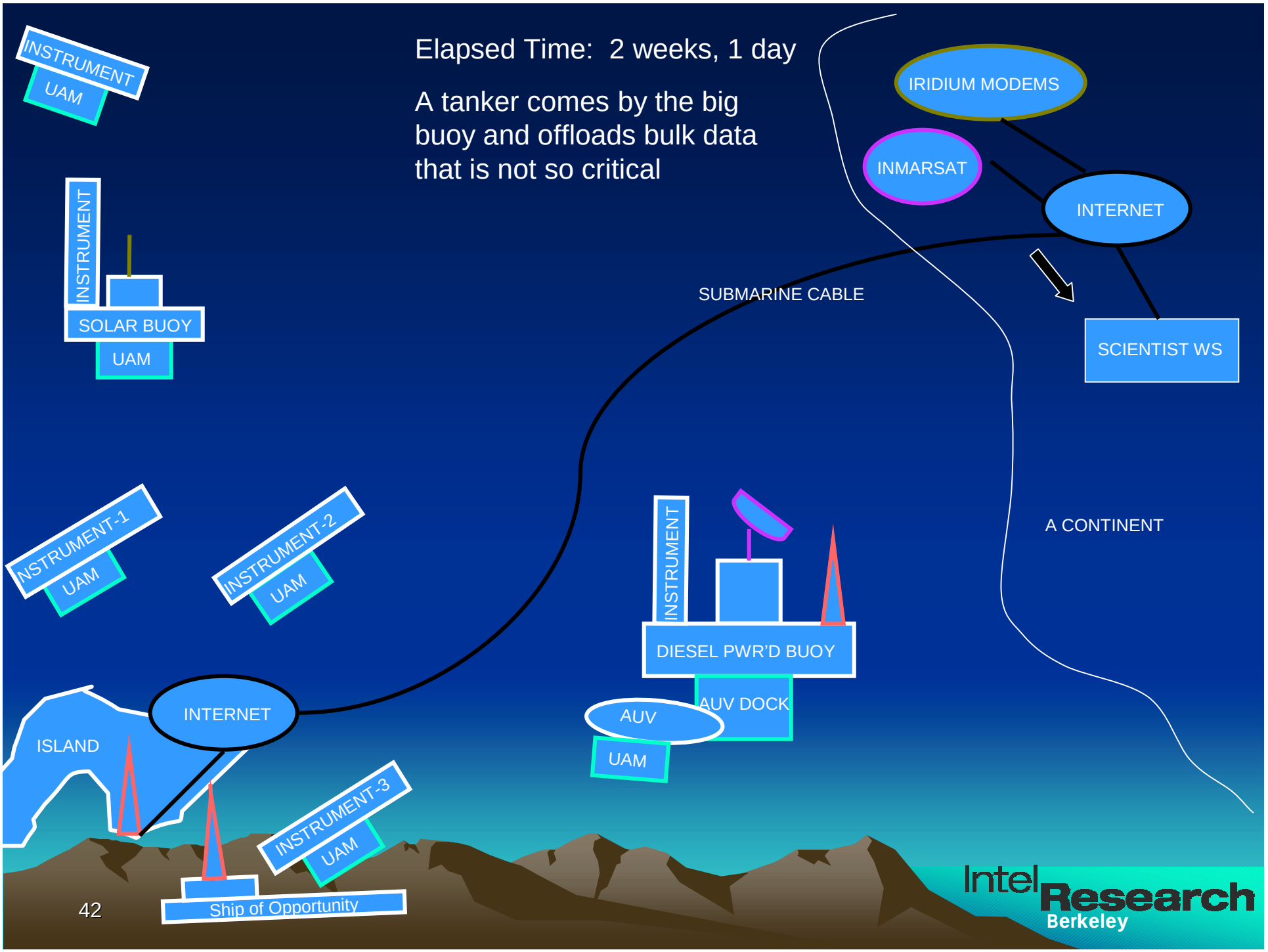
Intel Research Berkeley



Intel Research Berkeley



Intel Research Berkeley



Intel Research Berkeley

Reference Implementation Overview

- Written primarily in C++
  - ~22,000 non-comment lines of C++ (~5,000 in C)
  - 10K in generic system support classes (oasys)
  - 189 individual classes
  - Multithreaded using pthreads, mutex, spin lock
  - STL data structures (string, list, hashtable, …)
- Emphasis on clarity, cleanliness, flexibility
- Ported to Linux, Solaris, Win32 / Cygwin, Linux on PDA (ARM), FreeBSD, Mac OSX

(Source line statistics generated using David A. Wheeler's 'SLOCCount')

Status

- IETF/IRTF DTNRG formed end of 2002
  - See http://www.dtnrg.org
- DTN1 Agent Source code released 3/2003
- SIGCOMM Papers: 2003 [arch], 2004 [routing]
- Several other documents (currently ID's):
  - DTNRG Architecture document
  - Bundle specification
  - Application of DTN in the IPN
- Basis for new DARPA DTN program
- Part of NSF 'ICT4B' Project (with UCB)

Naming and Addressing

- Support ‘radical heterogeneity’ using *regions*:
  - Regions define a namespace (or address space)
  - *May* be defined based upon network topology
- Endpoint Name: ordered name pair **{R,A}**
  - **R**: region name [globally valid], used as routing hint
  - **A**: admin ID-- region-specific, opaque outside region **R**
  - **example: {sys1.iridium, +18455551212}**
  - represent as an Internet-style URI [see RFC2396]
- **Late binding** of **A** helps isolated nodes:
  - Only resolve A to address [if necessary] in transit
    - A interpreted only by nodes assigned region R
    - semantics implemented only in appropriate region

Acknowledgements

- DTN/ICT4B Contributors:
  - Eric Brewer, Mike Demmer, Rabin Patra (UCB)
  - Bob Durst, Keith Scott (MITRE)
  - Kevin Fall, Melissa Ho (Intel Research Berkeley)
  - Sushant Jain (Univ. of Washington)
  - S. Keshav (U Waterloo)
  - Ting Liu (Princeton)
  - Vint Cerf (MCI)
  - Scott Burleigh, Adrian Hooke (NASA/JPL)
  - Stephen Farrell (Trinity College, Ireland)
  - The *dtn-interest* mailing list and DARPA
- Thanks also to Andy Maffei & Matt Grund (WHOI)

Matt Grund's [whoi] q's

- How do you buffer the data?
  - in files or in database system [abstract api]
- How do you decide when to send the data?
  - Depends on type of link:
    - if 'on demand', then right away
    - if 'scheduled', then not until schedule tells us to [also, some notion of remote side initiating connection– NAT]
- What's the Ack scheme? How success indicated?
  - hop-by-hop acks ~ "custody transfer"
  - end-to-end "return receipt", if requested
- Could you demo it at WHOI?
  - yes, but not now ☺